

国家出版基金项目
NATIONAL PUBLICATION FOUNDATION

中外物理学精品书系

前沿系列·15

硅基光电子学

〔美〕周治平 主编

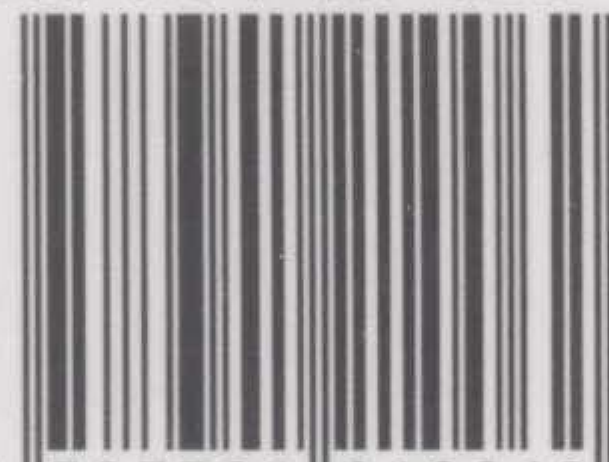


北京大学出版社
PEKING UNIVERSITY PRESS

中外物理学精品书系·前沿系列15

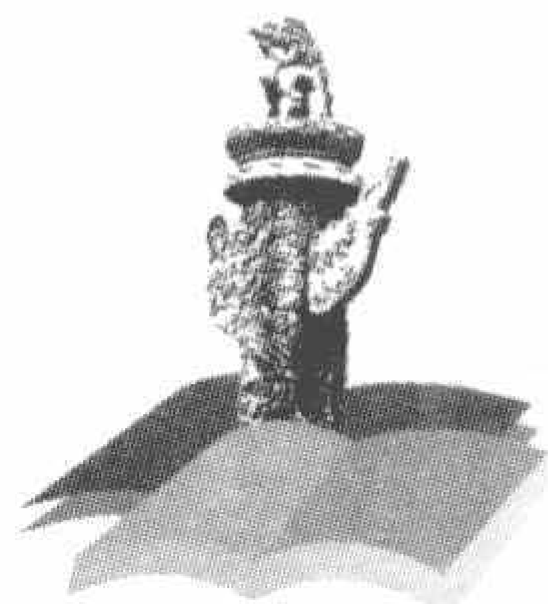
硅基光电子学

周治平 北京大学信息科学技术学院教授，长江学者特聘教授，SPIE Fellow，IET Fellow，中国光学学会第六、第七届理事会理事，IEEE中国武汉分会创会主席（2006—2008）。1993年获美国乔治亚理工学院博士学位。长期从事微纳光电子器件及其集成技术的研究和开发，在硅基光源、纳米光栅、光电调制、光学传感、光子晶体、表面等离子体器件及硅基集成技术等方面有独创性成果。发表论文、特邀报告270余篇，参编书籍5部，拥有国际国内专利15项。讲授过现代电子通信、集成电路、光电子学、硅基光电子学等课程。



ISBN 978-7-301-21006-2

定价：54.00元



中外物理学精品书系

前沿系列·15

硅基光电子学

〔美〕周治平 主编


北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

硅基光电子学/(美)周治平主编. —北京：北京大学出版社，2012.8
(中外物理学精品书系·前沿系列)
ISBN 978-7-301-21006-2

Ⅰ.①硅…　Ⅱ.①周…　Ⅲ.①硅基材料—光电子学—研究　Ⅳ.①TN304.2

中国版本图书馆 CIP 数据核字(2012)第 166412 号

书　　名：硅基光电子学
著作责任者：〔美〕周治平　主编
责 任 编 辑：王剑飞
标 准 书 号：ISBN 978-7-301-21006-2/O·0877
出 版 发 行：北京大学出版社
地　　址：北京市海淀区成府路 205 号　100871
网　　址：http：//www.pup.cn　　电子信箱：zpup@pup.pku.edu.cn
电　　话：邮购部 62752015　发行部 62750672　编辑部 62765014
出版部 62754962
印　刷　者：北京中科印刷有限公司
经　销　者：新华书店
730 毫米×980 毫米　16 开本　19.75 印张　357 千字
2012 年 8 月第 1 版　2012 年 8 月第 1 次印刷
定　　价：54.00 元

中外物理学精品书系

本书出版得到“国家出版基金”资助

《中外物理学精品书系》
编 委 会

序　　言

物理学是研究物质、能量以及它们之间相互作用的科学。她不仅是化学、生命、材料、信息、能源和环境等相关学科的基础,同时还是许多新兴学科和交叉学科的前沿。在科技发展日新月异和国际竞争日趋激烈的今天,物理学不仅囿于基础科学和技术应用研究的范畴,而且在社会发展与人类进步的历史进程中发挥着越来越关键的作用。

我们欣喜地看到,改革开放三十多年来,随着中国政治、经济、教育、文化等领域各项事业的持续稳定发展,我国物理学取得了跨越式的进步,做出了很多为世界瞩目的研究成果。今日的中国物理正在经历一个历史上少有的黄金时代。

在我国物理学科快速发展的背景下,近年来物理学相关书籍也呈现百花齐放的良好态势,在知识传承、学术交流、人才培养等方面发挥着无可替代的作用。从另一方面看,尽管国内各出版社相继推出了一些质量很高的物理教材和图书,但系统总结物理学各门类知识和发展,深入浅出地介绍其与现代科学技术之间的渊源,并针对不同层次的读者提供有价值的教材和研究参考,仍是我国科学传播与出版界面临的一个极富挑战性的课题。

为有力推动我国物理学研究、加快相关学科的建设与发展,特别是展现近年来中国物理学者的研究水平和成果,北京大学出版社在国家出版基金的支持下推出了《中外物理学精品书系》,试图对以上难题进行大胆的尝试和探索。该书系编委会集结了数十位来自内地和香港顶尖高校及科研院所的知名专家学者。他们都是目前该领域十分活跃的专家,确保了整套丛书的权威性和前瞻性。

这套书系内容丰富,涵盖面广,可读性强,其中既有对我国传统物理学发展的梳理和总结,也有对正在蓬勃发展的物理学前沿的全面展示;既引进和介绍了世界物理学研究的发展动态,也面向国际主流领域传播中国物理的优秀专著。可以说,《中外物理学精品书系》力图完整呈现近现代世界和中国物理科学发展的全貌,是一部目前国内为数不多的兼具学术价值和阅读乐趣的经典物理丛书。

《中外物理学精品书系》另一个突出特点是,在把西方物理的精华要义“请进来”的同时,也将我国近现代物理的优秀成果“送出去”。物理学科在世界范围内的重要性不言而喻,引进和翻译世界物理的经典著作和前沿动态,可以满足当前国内物理教学和科研工作的迫切需求。另一方面,改革开放几十年来,我国的物理学研究取得了长足发展,一大批具有较高学术价值的著作相继问世。这套丛书首次将一些中国物理学者的优秀论著以英文版的形式直接推向国际相关研究

的主流领域，使世界对中国物理学的过去和现状有更多的深入了解，不仅充分展示出中国物理学研究和积累的“硬实力”，也向世界主动传播我国科技文化领域不断创新的“软实力”，对全面提升中国科学、教育和文化领域的国际形象起到重要的促进作用。

值得一提的是，《中外物理学精品书系》还对中国近现代物理学科的经典著作进行了全面收录。20世纪以来，中国物理界诞生了很多经典作品，但当时大都分散出版，如今很多代表性的作品已经淹没在浩瀚的图书海洋中，读者们对这些论著也都是“只闻其声，未见其真”。该书系的编者们在这方面下了很大工夫，对中国物理学科不同时期、不同分支的经典著作进行了系统的整理和收录。这项工作具有非常重要的学术意义和社会价值，不仅可以很好地保护和传承我国物理学的经典文献，充分发挥其应有的传世育人的作用，更能使广大物理学人和青年学子切身体会我国物理学研究的发展脉络和优良传统，真正领悟到老一辈科学家严谨求实、追求卓越、博大精深的治学之美。

温家宝总理在2006年中国科学技术大会上指出，“加强基础研究是提升国家创新能力、积累智力资本的重要途径，是我国跻身世界科技强国的必要条件”。中国的发展在于创新，而基础研究正是一切创新的根本和源泉。我相信，这套《中外物理学精品书系》的出版，不仅可以使所有热爱和研究物理学的人们从中获取思维的启迪、智力的挑战和阅读的乐趣，也将进一步推动其他相关基础科学更好更快地发展，为我国今后的科技创新和社会进步做出应有的贡献。

《中外物理学精品书系》编委会　主任

中国科学院院士，北京大学教授

王恩哥

2010年5月于燕园

内容简介

本书是笔者在微纳光电子领域多年研究和教学基础上完成的，系统描述了硅基光电子学的基础理论、器件原理及应用前景。全书共10章。第1～3章讲述了硅基光电子学的起源及所需的基本知识；第4章介绍了硅基无源器件；第5～8章为硅基有源器件，包括光源、调制器、探测器、表面等离子体激元器件等；第9章介绍器件工艺与系统集成；第10章探讨了硅基光电子学的应用。

本书可作为高等院校电子学、光电子学、物理电子学、微电子与固体电子学、通信与信息系统、计算机技术等专业高年级本科生和研究生相关课程的教材。对于在相关领域内工作的研究人员和工程技术人员，本书也是一本有用的参考书。

前　言

晶体管这个被誉为20世纪最伟大的发明改变了世界，而以硅材料为基础的微电子器件则以其低功耗、低成本、易集成等优点迅速占领了绝大部分电子市场，并成为当时高科技产业的重要支柱. 从晶体管的发明、集成电路的出现，到计算机的不断更新换代，再到通信网络的飞速发展，人类生活的各个层面无不打下了微电子的烙印——微电子产品已经被用到人们物质生活的各个层面；微电子的设计思想和制作方法也已经渗入到不同学科及领域.

然而，微电子器件的进一步小型化使得集成电路的互连延迟效应及能耗问题成了以电子作为信息载体的高速集成电路技术的一个不可逾越的障碍. 与电子相比，光子作为信息载体具有巨大的优势：光子没有静止质量，光子之间也几乎没有干扰，光的不同波长可用于多路同时通信，因此利用光子的通信信息技术具有更大的带宽和更高的速率.

所谓硅基光电子学，就是研究、开发以光子和电子为信息载体的硅基大规模集成技术. 其核心内容就是研究如何将光子器件"小型化"、"硅片化"并与纳米电子器件相集成，即利用硅或与硅兼容的其他材料，应用硅工艺，在同一硅衬底上同时制作若干微纳量级，以光子、电子为载体的信息功能器件，形成一个完整的具有综合功能的新型大规模集成芯片.

虽然硅材料在光电效应方面存在着"先天不足"，而光子器件在尺寸方面也"衍射受限"，但将微电子和光电子结合起来，开发硅基大规模光电子集成技术，已经成为信息技术发展的必然和业界的普遍共识. 本书的目的之一在于为这一历史进程提供理论研究和技术开发的支持，希望其内容能够对读者起到一个"抛砖引玉"的作用.

本书是笔者在微纳光电子领域多年研究和教学基础上完成的，系统描述了硅基光电子学的基础理论、器件原理及应用前景. 全书共10章. 第1～3章讲述了硅基光电子学的起源及所需的基本知识；第4章介绍了硅基无源器件；第5～8章为硅基有源器件，包括光源、调制器、探测器、表面等离子体激元器件等；第9章介绍器件工艺与系统集成；第10章探讨了硅基光电子学的应用. 笔者及其所领导的硅基光电子及微系统研究团队是本书编写过程中的主要贡献者. 其中笔者编写了第1,2,6章，杨俊波编写了第3章，吴华明编写了第4章，王兴军编写了第5,7章，肖功利编写了第8章，冯俊波编写了第9章，易华祥编写了第10章. 在成书过程中，王雨竹、侯金、王冰、郭瑞民、胡飞飞、王磊、伍丹华、高琳斐、谈

卫、龙祁峰、涂芝娟、尹兵、尹美、陈特、臧凯、袁格、洪凌宇、王人泽等提供了相关协助及修改意见.

感谢科技部、国家自然科学基金委员会、北京大学、华中科技大学、乔治亚理工学院为本书的编写在科研项目和工作环境方面的支持. 也感谢 IEEE Ⅳ族光电子学国际会议(International Conference on Group Ⅳ Photonics)这一信息交流平台为本书的形成所提供的部分思路和灵感.

周治平

2011 年 12 月 24 日于北京

目　　录

第一章　绪　　论

§1.1　从微电子到光电子

微电子时代的发展可以追溯到晶体管的出现. 1947 年 12 月 16 日，美国 AT&T 公司 Bell 实验室的三位科学家 William Shockley, John Bardeen 和 Walter Brattain 制成了世界上第一支晶体管，开始了以晶体管代替电子管的时代，从而拉开了微电子时代的序幕. 晶体管这个被誉为 20 世纪最伟大的发明改变了世界，而以硅(Si)材料为基础的半导体工业迅速占领了绝大部分市场，并成为当时高科技产业的重要支柱. 从晶体管的发明、集成电路(integrated circuit, IC)的出现、计算机的不断更新换代，再到通信网络的飞速发展，人类生活的各个层面无不打下了微电子的烙印——微电子产品已经被用到人们物质生活的各个层面；微电子的设计思想和制作方法也已经渗入到不同学科及社会领域. 这一时代被誉为“微电子时代”是恰如其分的. 但是，随着微电子技术往纵深方向发展，其量子特征也逐渐表现出来. 近 50 年来一直成功地描述了 IC 发展趋势的摩尔定律也正在接近极限.

1.1.1　微电子技术所面临的挑战

微电子技术是建立在以三极管等半导体器件为核心的集成电路基础上的，其主要产品集成电路芯片是构成各种电子电路和电子信息系统的核心部件. 60 多年来，微电子技术按照摩尔定律预测的那样，“半导体芯片的集成度每 18 个月增长一倍，而价格则降低一半”，以惊人的速度一直发展着. Intel 公司曾宣布，从 2011 年的下半年开始，将采用新的 22 nm 工艺生产新一代的微处理器[1]. 而根据国际半导体技术蓝图(ITRS)2009 年发布的半导体技术路线图预测，到 2016 年时，世界集成电路闪存线宽将减小至 15.9 nm；而到 2024 年时，最小线宽更是减小达到 6.3 nm，接近理论极限[2]. 届时，以硅为基础的微电子技术由于物理极限的限制，将很难继续遵循摩尔定律来发展[1]. 首先，随着微纳器件集成度的进一步提高，器件线宽的进一步减小，电互连所固有的局限性将促使芯片的发热量迅速增加，引起串扰、噪声、功耗、时延等多方面的问题，从而使器件无法正常工作；其次，是现有的加工设备已经接近工艺极限，通过减小线宽的方法来提高芯片的工作频率和集成度面临非常大的工艺问题；此外，当线宽进入深纳米尺寸时，如何避免量子效应导致相邻导线之间的量子隧穿，也面临前所未有的挑战[3~7].

可见,用电子作为信息载体的微电子技术,当器件的加工线宽发展到深纳米尺度时,将遇到其发展的理论与技术瓶颈.因此,亟需新的理论与技术解决上述诸多限制与问题,否则信息技术的发展将裹足不前,并严重影响人类社会和经济的进一步向前发展[3~7].

1.1.2　集成光路的发展

在21世纪,人们将迈入一个高度信息化的时代.信息时代的特征是:信息量十分巨大,信息传递非常快捷,信息处理十分迅速.目前大部分的信息是由电子传导的.电子具有静止质量,且电子之间存在库仑作用,因此电子的运动及电信号易受电磁场干扰,从而限制了电子通信的容量和速率.与电子相比,光子作为信息载体具有巨大优势:光子没有静止质量,光子之间也几乎没有干扰,光的不同波长可用于多路同时通信,因而光信息传输可具有更大的带宽和更高的速率.另一方面,光信号处理速度很高,且不受电磁场干扰.这些优点使得光子技术在未来的信息化社会中必将扮演非常重要的角色.

受到了微电子集成电路技术的启迪和促进,光子学发展的一个重要目标就是要实现集成光路[8].集成光路是利用光波导将光发射、光耦合、光传输、光调制、光逻辑、光处理以及光接收等器件连接在一起,集成到同一基底上,构成的具有一定独立功能的微型光子学系统.

图1.1为一套典型的集成光路系统结构示意图,它包含了光的产生、耦合、传输、调制、探测等几个部分.

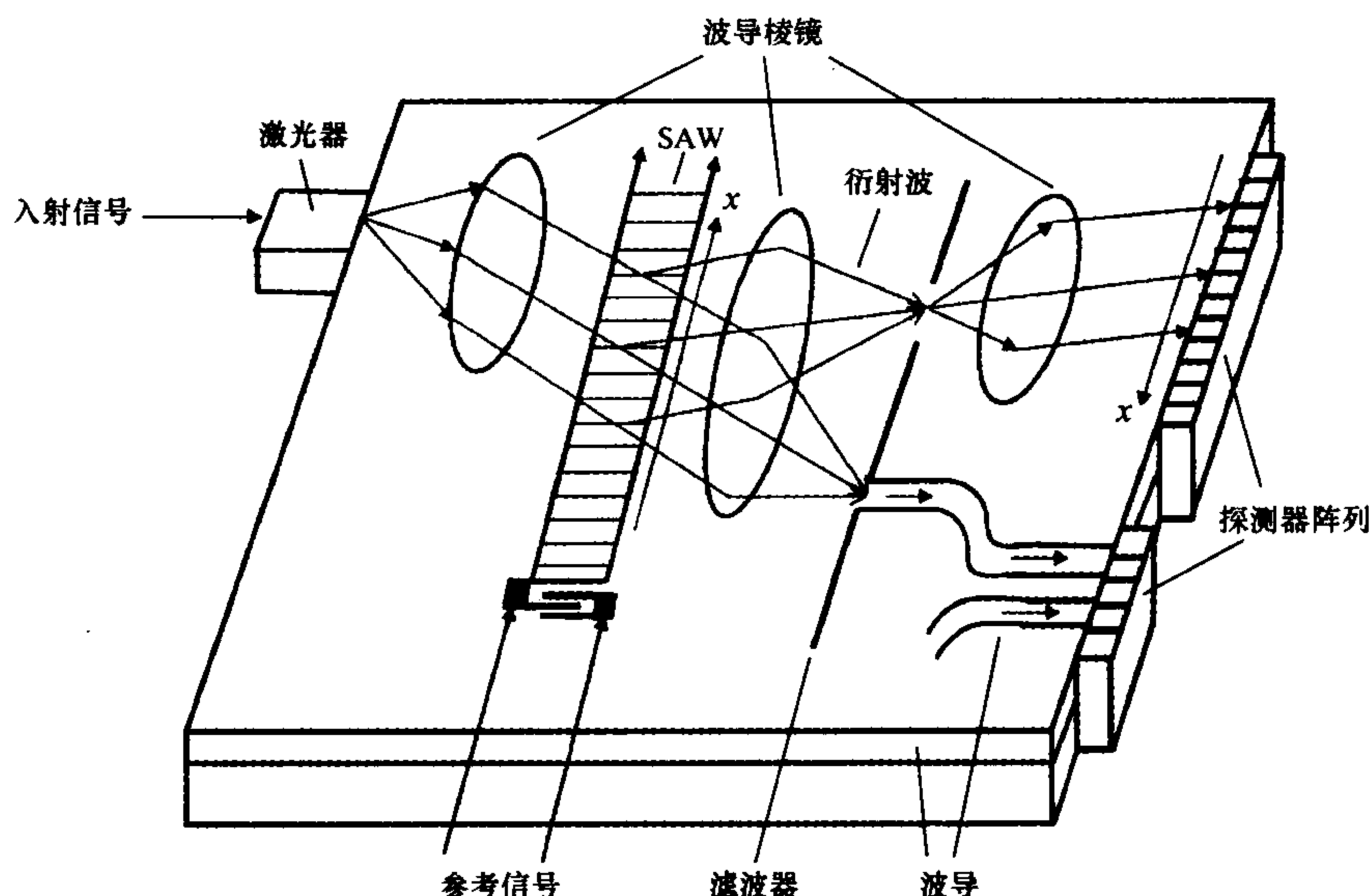


图1.1　一套典型的集成光路系统结构示意图

目前,集成光路的发展已经初具规模,并在光通信及光信息处理方面显示出电子学无法比拟的优越性.与普通光学系统相比,集成光路具有很多优点:信号带宽大,容易实现密集波分复用,尺寸小,重量轻,功耗小,成批制备成本低,可靠性高等.得益于微电子制备技术和光电子学的发展,集成光路正以其独特的优势进入了迅速发展的阶段.表1.1列出了集成光路和集成电路的比较.

表1.1　集成光路和集成电路的比较

	集成光路	集成电路
基本作用	光波导中的光传输及光与电子、晶格的相互作用	基底表面附近的电子传输与控制
基本元器件	光波导、半导体激光器、探测器等	晶体管、电阻、电容等
元器件尺寸(厚度方向)	波长量级、纳米量级	数十纳米至微米
元器件尺寸(长度方向)	微米至数毫米,关键部分甚至纳米量级	数十纳米至微米
器件之间的连接	光波导,难点在于需要精确的耦合对准	电气布线,优点在于直接导体连接
元器件可靠性	目前还存在问题,通常需要检测全部元件	几乎没问题,通常只进行抽检
元器件制作工艺	多样,目前还在研究开发中	成熟的工艺

但是,集成光路目前在实践应用中仍然面临许多尚未克服的技术和性能障碍,在理论上也有待新的突破.最主要的问题是集成光路中还没有统一的标准化制备技术,导致制造成本很高.这主要是由于集成光路和微电子制备技术相比,选用的材料、各个元器件的制备工艺和工作原理等有很大不同,如集成光路选用的材料有Si,$LiNbO_3$,InP等,微电子制备技术的基体材料是Si.微电子中的基本单元是晶体管,而集成光路的基本单元包括激光器、探测器、调制器、滤波器等,各个器件的工作原理都不相同.当前,集成光学的研究热点主要集中在以下几个方面:

① 提高集成光路的可靠性和稳定性;

② 减小器件尺寸,提高集成度;

③ 制备工艺优化,减小光波导传输损耗;

④ 突破目前的“光-电-光”转换瓶颈;

⑤ 提高光通信用多波长光源的性能、成本、可靠性;

⑥ 研究新的耦合技术以便实现高性能的耦合,解决光纤或激光器与波导的耦合难题.

然而，由于光逻辑和光处理器件的发展相对缓慢，具有真正“计算”功能的集成光路还正处于早期探索阶段.

1.1.3 光电子集成

众所周知，光和物质间存在着诸如光电、电光、声光、磁光，热光等多种线性或非线性物理效应，可以用来实现各种功能丰富的微纳光电器件. 当纯粹的电子或光子集成遇到困难时，将两者结合不失为一种折中的选择. 特别是，在集成电路和集成光路都有了相当程度积累的今天，努力发展光电子集成就显得非常自然了. 可以预见，未来的信息化时代将是光子载体与电子载体携手合作，互相补充，最终走向融合统一，实现光电集成回路的时代[3~7].

图 1.2 为 20 世纪 90 年代由 R. A. Soref 建议的硅基光电子混合集成芯片[9]，包含了光的产生、放大、耦合、传输、调制、探测和高速电子电路等部分. 其光信号通过光纤与波导直接耦合的方式与外界相连. 所有的光电子器件均由Ⅲ-Ⅴ族半导体构成，并且与电子电路分别制作，混合集成.

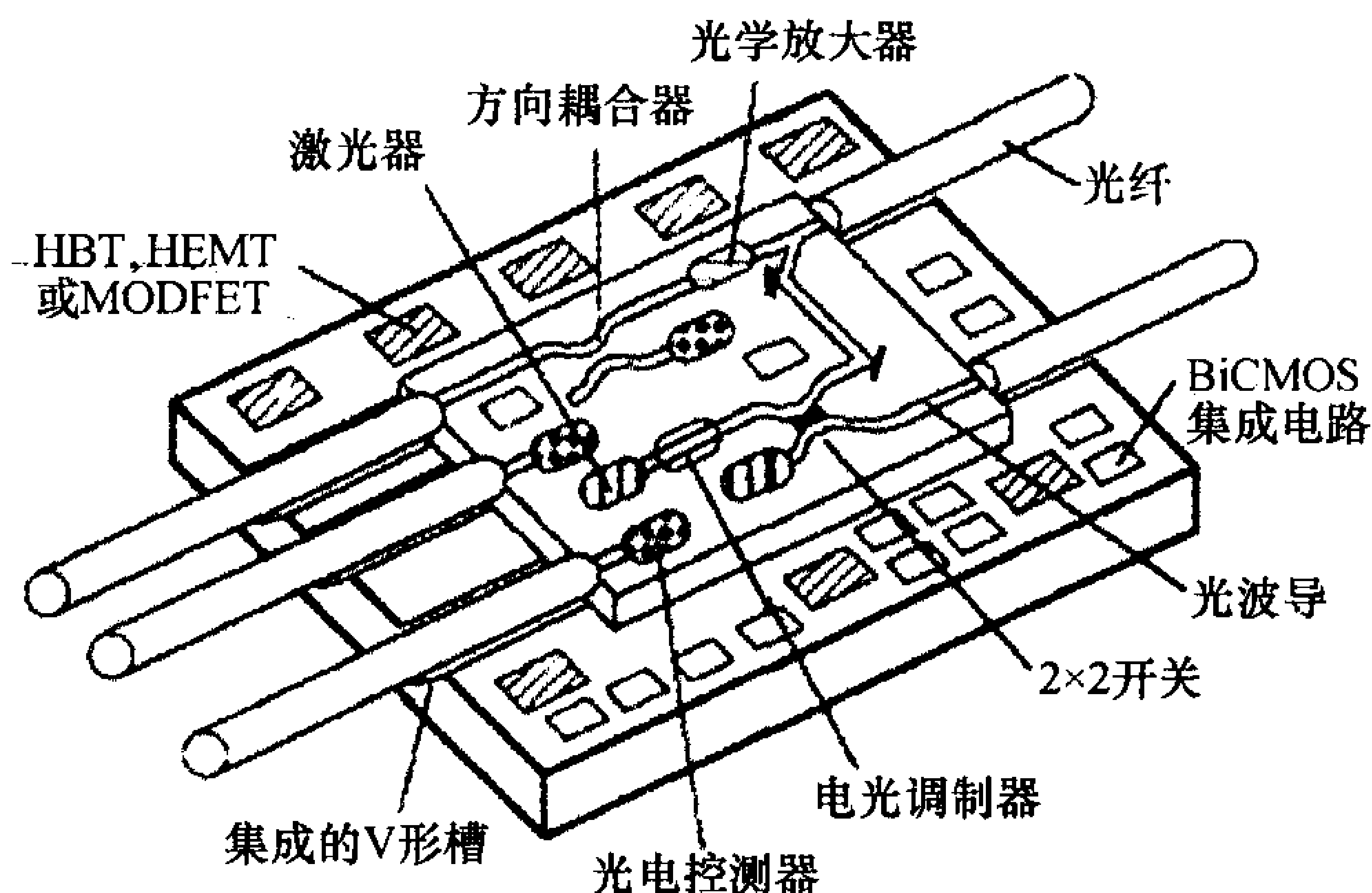


图 1.2 20 世纪 90 年代的硅基光电子混合集成系统示意图

图 1.3 是 Intel 公司给出的光互连与电互连（主要指铜线互连）在成本和带宽关系上的博弈图[10]. 从图中我们可以看出，在当前以及今后一段时期内光互连技术与电互连技术将共同存在，但是当芯片的工作频率扩展到 40 GHz 以上时，光互连的优势将凸显，并且将逐渐取代电互连成为芯片互连技术的主体.

然而，应用于以电子为主要信息载体的集成芯片中的光互连技术是一定要与电子及电子技术相互作用的. 光子技术与电子技术在硅基底上的融合导致了硅基光电子学的发展.

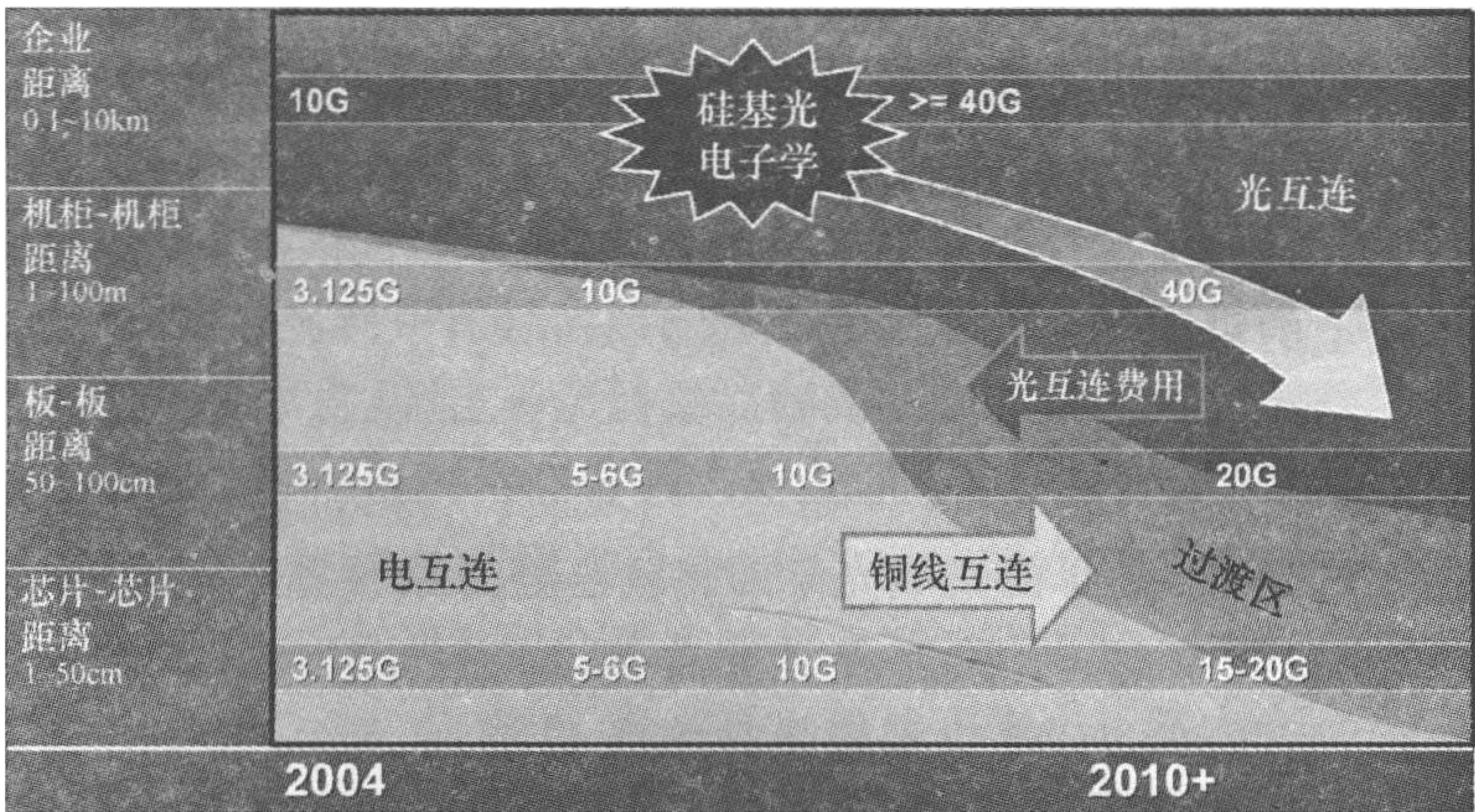


图 1.3 光互连与电互连的成本、带宽、传输距离关系以及发展趋势图

§1.2 硅基光电子学的发展

1.2.1 硅基光电子学

所谓硅基光电子学，就是研究如何利用硅或与硅工艺兼容的其他材料，开发以光子、电子为载体的微纳量级信息功能器件，并将它们在同一硅基底上大规模集成，形成一个完整的具有综合功能的新型芯片单元. 学科所涉及的领域包括衍射和纳米光学、超快速光通信、表面等离子体理论与技术、半导体器件物理、半导体器件工艺、材料能带工程、集成传感技术、纳米科学与技术、微纳加工及集成技术等.

之所以选用硅作为光电子器件的材料平台，是因为以硅为主导的微电子技术在过去半个世纪中取得了举世瞩目的成就，大力推动了信息技术的发展. 目前，人们对硅及其衍生物各种特性的了解，已经居自然界所有元素之首. 而且，全世界数以万亿美元的设备和技术投入，以及几十年的技术积累，使得硅基工艺具有非常强大的产业能力：根据 2003 年末的报道，硅占据全世界 97.2%的半导体材料市场，而 III-V 族等其他半导体仅占 2.8%左右的份额[11]. 硅材料具有如下优势：

① 硅片尺寸大，机械性能好，加工方便，相比其他材料成本更低；

② 折射率大，具有良好的波导特性；

③ 通信用波长传输透明；

④ 硅基光电子工艺与先进的互补金属氧化物半导体（complementary metal

oxide semiconductor，CMOS)工艺兼容，具有丰厚的工艺技术积累；

⑤ 绝缘体上硅(silicon on insulator，SOI)在光学上具有很好的导光性质，在电学上具有很好的抗辐射性能.

以上这些因素，使得硅在光电子学时代占据主导作用不仅具有理论上的依据，更有实践上的可能.

1.2.2 硅基光电子学的发展

20世纪70年代，最早的集成光学回路就已经开始被提出，但是早期的工作都是基于铁电物质 $LiNbO_3$ 和Ⅲ-Ⅴ族化合物. $LiNbO_3$ 具有很大的电光系数可以通过泡克耳斯(Pockels)效应实现快速的光调制，而Ⅲ-Ⅴ族化合物在实现激光器、光放大器等方面具有非常好的性能，一直以来它们都吸引着众多研究者极大的兴趣. 尽管如此，硅材料在半导体微电子应用中的主导作用以及其在集成方面的潜在优势，最终引发了研究者开展基于硅基光电子集成回路的研究工作，从20世纪80年代中期至今一直持续并发展着，开始的时候并没有受到很大的关注. 这是由于硅为间接带隙半导体，载流子直接跃迁复合的发光效率很低，这意味着不可能用常规手段在硅上作出高效率的发光器件；其次，硅具有中心反演对称的晶格结构，不存在直接电光效应，其载流子迁移率比 GaAs 中的低得多，限制了硅基光电子器件向更高的工作频率和速度发展. 正是由于这两大缺点使得人们最初并不看好硅在光学和光子学领域的前景，硅基光电子学并没有得到什么发展. 但近些年，一些硅的新物理现象的发现引发了人们研究硅基光电子学的热潮.

1987年，Soref 和 Bennett[12]研究发现材料中自由载流子的注入在调制机制上具有潜在的应用，通过改变载流子的浓度能改变硅的折射率，这就是等离子体色散效应. 由于硅材料不具有线性光电效应，所以一般硅基光调制器和光开关是基于硅的热光效应和等离子色散效应而设计的. 硅的热光效应是指其折射率随材料温度的高低而变化，而等离子色散效应是指硅的折射率和对光的吸收系数随材料中载流子浓度而变化. 当利用热光效应对光进行调制时，由于牵涉到热的扩散过程，其调制速率较低，因而调制频率很难达到1 MHz以上，故高速率的调制器件一般采用等离子体色散效应进行工作. 近几年来，硅基光调制器的研究进展令人鼓舞. 2004年 Intel 公司在《自然》(*Nature*)杂志上报道了调制带宽超过1 GHz的硅基高速光波导调制器[13]，它将器件的调制方式从常规的电流调制转化为电压调制，从而实现对光的相位进行快速调制. Mao 等[14]设计了双 MOS 结构和 n-p-n 结构硅基电光调制器，其调制速率分别达到15 GHz和14 GHz. 2009年，Lipson 等[15]使用微环波分复用结构实现了50 GHz的调制速率.

由于硅的载流子直接跃迁复合的发光效率很低，因此用常规手段在硅上作

出高效率的发光器件比较困难. 虽然目前已经有一些单片集成硅基光源的成果报道,但是在短期内真正实现高效的单片集成硅基激光仍具有挑战性. 目前使用的方式有外部光源和混合集成的片上光源两种. 其中,外部光源又分为芯片外部光源和倒装焊外部光源两种,如图 1.4 所示. 外部光源的方案回避了直接在硅材料上制备光源的难题,利用比较成熟的外部光源作为系统的输入. 但这种方案使得系统无法大规模集成,而且必须解决外部光源到集成光电子器件的耦合问题. 对于倒装焊外部光源需要严格和苛刻的对准条件,而混合集成的片上光源的结构紧凑. 一种变通的方法是利用键合等非 CMOS 兼容工艺将Ⅲ-Ⅴ族化合物集成在同一基底上. 外部光源和混合集成的片上光源具有较高的发光效率,容易实现连续或者可调激光输出的优点,但其缺点仍然很突出: 集成度不高,制备工艺复杂,与 CMOS 工艺不兼容,系统成本昂贵,而且必须考虑到输出激光到集成光电子器件的耦合对准问题. 因此,将光源单片集成到硅基光电子系统仍然是当前研究的重中之重.

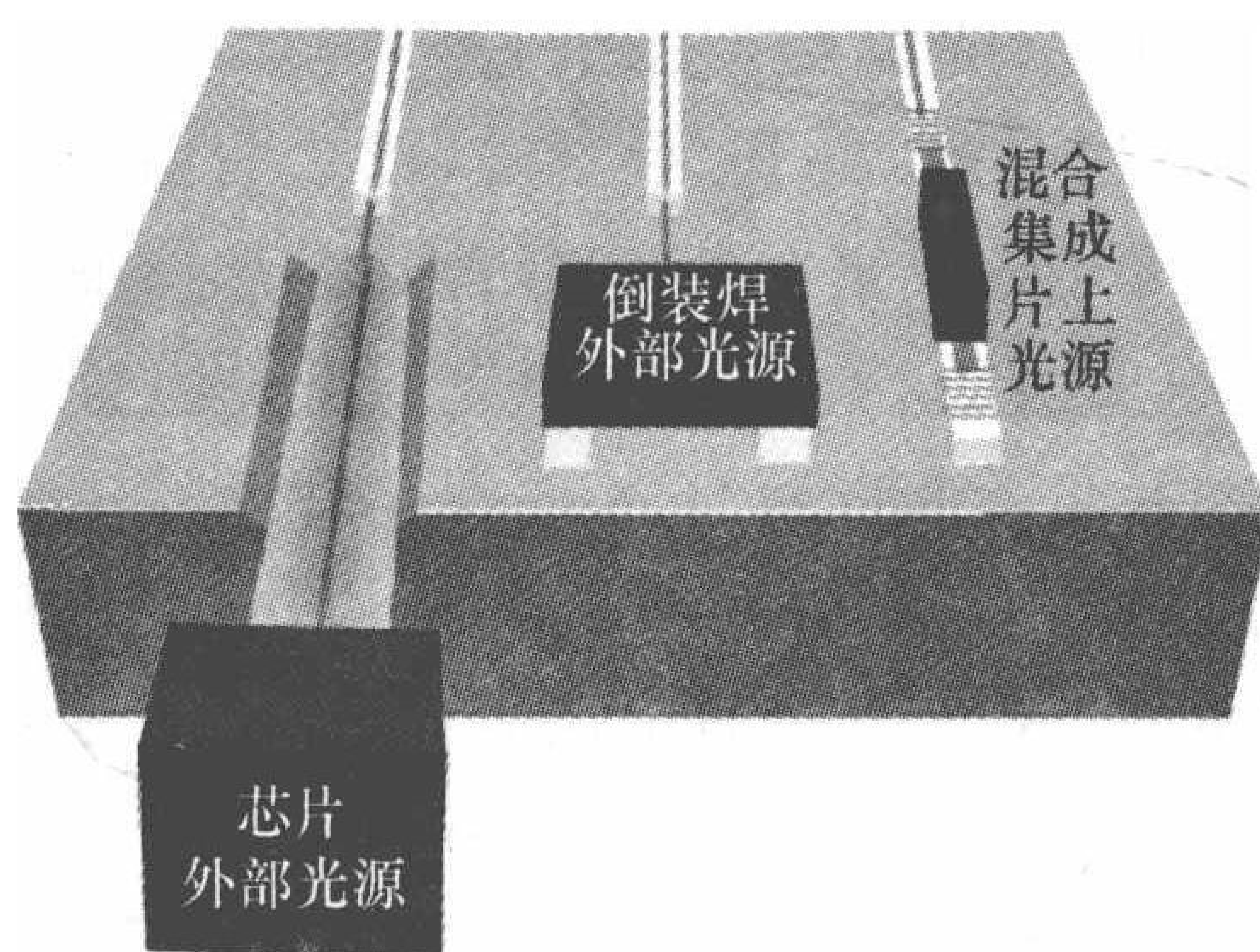


图 1.4　外部光源及混合集成的片上光源示意图

自 2000 年以来,一系列优秀研究成果逐渐出现,如体硅发光二极管(light emitting diode, LED)、富硅氧化硅发光等. 美国 Intel 公司的研究人员利用由多层介电薄膜覆盖的 SOI 光波导形成的光学微腔,实现了稳定的单模激射输出,并研制成了连续拉曼激光器,从而成为硅基激光器发展史上的一个重要里程碑[16]. 掺铒(Er)硅和硅外延锗(Ge)光源是另外两个重要的研究方向. 其中,铒的硅酸盐化合物由于高的铒离子浓度(10^{22} cm^{-3},比其他方法高两个数量级),在较小的温度消光情况下获得了较强的室温光致发光,而引起了国际上的广泛关注[17]. Wang 等[18,19]采用四能级模型,通过铒速率方程和传输方程数值计算了铒的硅酸盐化合物光波导放大器光增益与铒浓度、波导长度、泵浦功率及信号功率的关系,当铒浓度为 0.5×10^{22} cm^{-3},泵浦功率为 70 mW 时,在 1 mm 波导长度下可以获得 10 dB 以上的增益. 锗的直接带隙只比间接带隙高 136 meV,早在

1978 年人们就观察到了其在低温下由高功率密度激光激发产生的直接带隙发光. 近年来,美国麻省理工学院 Kimerling 研究组[20]提出了通过能带工程获得外延锗光源的理论,并获得了令人瞩目的实验结果. 他们通过引入应变硅及重掺杂的方法,使 GeSi 变成直接带隙的半导体,大大提高了发光效率,并于 2010 年观察到了通信波段的激光输出.

而另一非常重要的有源器件硅基探测器用于光通信和光互连时所面临的问题,从本质上讲有两点:① 硅材料在波长大于 1100 nm 后即为透明,失去探测能力,如图 1.5 所示[21];② 硅材料中载流子运动能力较低,导致器件难以实现高速反应. 目前主要是通过外延其他材料(主要是锗),来解决硅基探测器的问题. 锗相对于硅有 4.2%的晶格失配,这使得生长锗层比较困难,容易在锗中产生高密度错位,因此如何生长高质量的锗层成为最重要的研究内容. 近年来,国际上不少研究机构在结构设计和材料生长上取得了很大的进展,其中具有代表性的有 IBM 公司和 MIT 的研究小组. 2008 年,Kang 等[22]报道了 GeSi 雪崩光电探测器,其增益带宽达到 340 GHz,性能达到了目前商用的Ⅲ-Ⅴ族化合物雪崩光电探测器的水平.

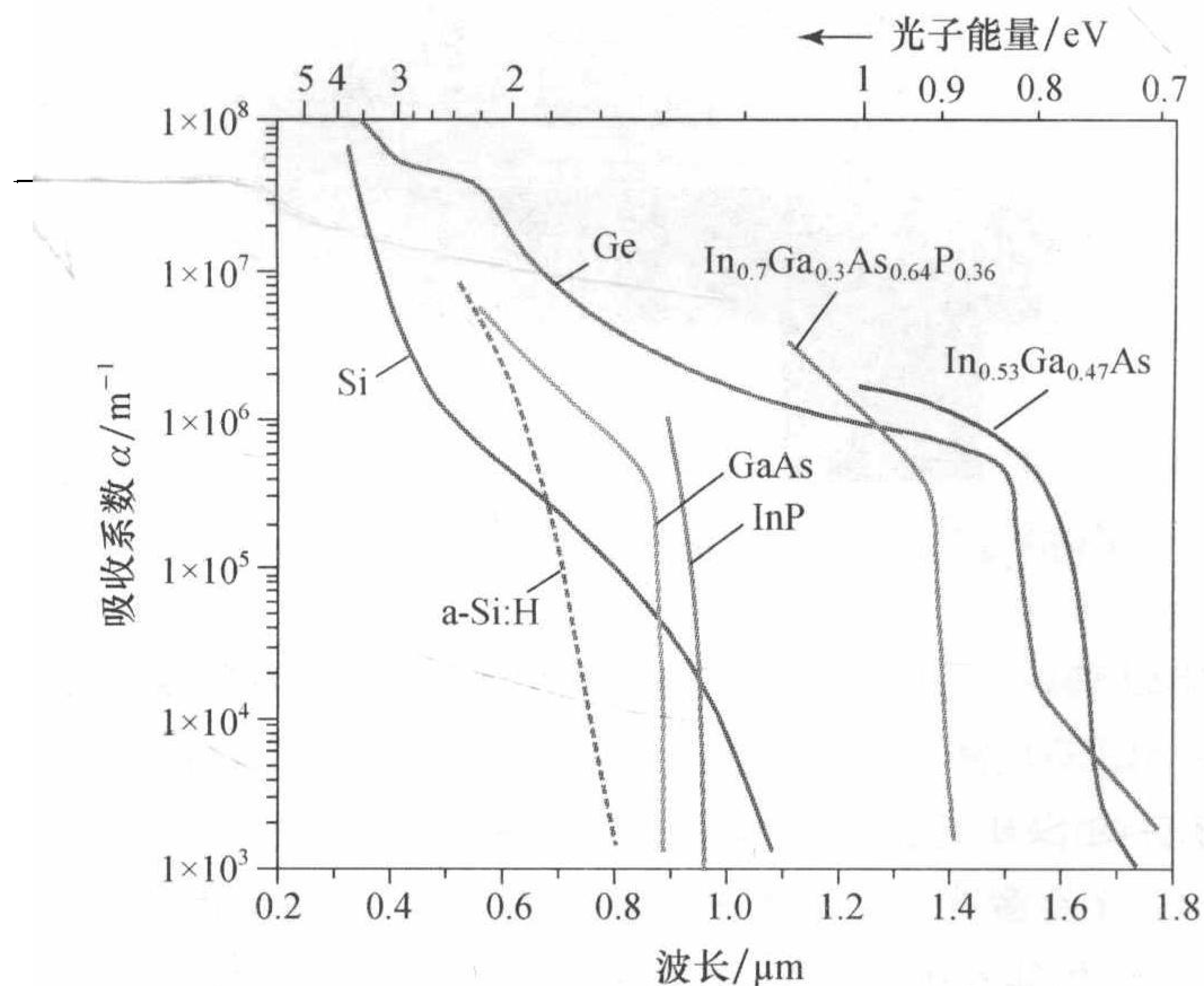


图 1.5 硅、Ⅲ-Ⅴ族材料吸收系数和波长的关系曲线

硅基光电子学的一个重要特征就是对以光子、电子为载体的信息功能器件都具有较高的集成度. 然而普通以光子为信息载体的介质波导结构的尺寸远大于微电子器件的尺寸,为减小器件尺寸,突破衍射极限的限制,对光子晶体(photonic crystal)和表面等离子体的研究应运而生. 1987 年,美国 Bell 实验室的

Yablonovitch[23]和普林斯顿大学的John[24]分别在研究如何抑制自发辐射和无序电介质材料中的光子局域时，各自独立地提出了光子晶体(photonic crystal, PC)的概念. 在过去的25年中，光子晶体获得了前所未有的关注和飞速的发展. 应用光子晶体的领域也在不断扩展，涉及光子集成、微波通信、光纤通信、太阳能发电 、生物化学传感和隐身技术等. 在光子集成领域，应用光子晶体的带隙特性，各种光发射器件、波导、分束器、耦合器、滤波器、光开关、调制器等逐渐从研究转向应用，其性能也不断提高. 近年来，光子晶体的慢光、自准直、负折射等非带隙特性成为了研究的新热点.

与光子晶体一样，表面等离子体激元(surface plasmon polariton, SPP)研究也正逐渐成为硅基光电子学领域一个热点问题. 表面等离子体是沿着导体表面传播的波，当改变金属表面结构时，表面等离子体激元的性质、色散关系、激发模式、耦合效应等都将产生重大的变化，通过与光场之间相互作用，能够实现对光传播的主动操控，它可以将光波长限制和控制在100 nm以下，是光子学和纳电子学的天然桥梁，能够实现电子线路的微型化和光传输. 最初，表面等离子体光子学主要集中在无源器件方面，如波导、开关、耦合器等[25,26]. 最近，基于Ⅲ-Ⅴ族和有机材料的有源等离子体器件也相继被报道. 2008年，Koller等[27]报道了基于有机半导体电致表面等离子体源. 2010年1月，Polman研究小组[28]又在*Nature Materials*报道了与CMOS技术兼容的硅基电致表面等离子体源. 表面等离子体的应用非常广泛，可以制成表面等离子体波导、光开关，以及利用近场增强效应改善发光效率，制造表面等离子体激光器等. Wang等[29,30]采用二维金属光子晶体来提高铒离子的发光效率. 由于金属光子晶体中的电磁场可以被局限在很小的范围内，例如在金属光子晶体的能带边缘处，表面等离子体的群速率接近零，因此相应的态密度较其他位置可提升数个量级. 通过将处于能带边缘处的表面等离子体耦合为光子并发射到远场，铒离子发光效率提高近40倍. 由于表面等离子体激元以极小的尺寸束缚在界面处，突破了衍射极限，可用于灵敏度很高的传感器结构. 此外由于表面的场强很高，也广泛用于非线性光学上的研究.

1.2.3 硅基光电子集成

在过去约60年的信息化进展中，一方面，以硅为基础的集成电路技术通过不断地提高"集成度"为科学技术的进步做出了卓越的贡献. 另一方面，以不同材料为平台的集成光学技术也在努力使光学器件"小型化"，从而催生了纳米光子器件. 在光子与电子相互融合、统一的发展过程中，人们自然希望能够继承当前硅基微电子产业成熟的理论与技术，在硅材料平台上面实现微纳光电子集成，如图1.6所示[31].

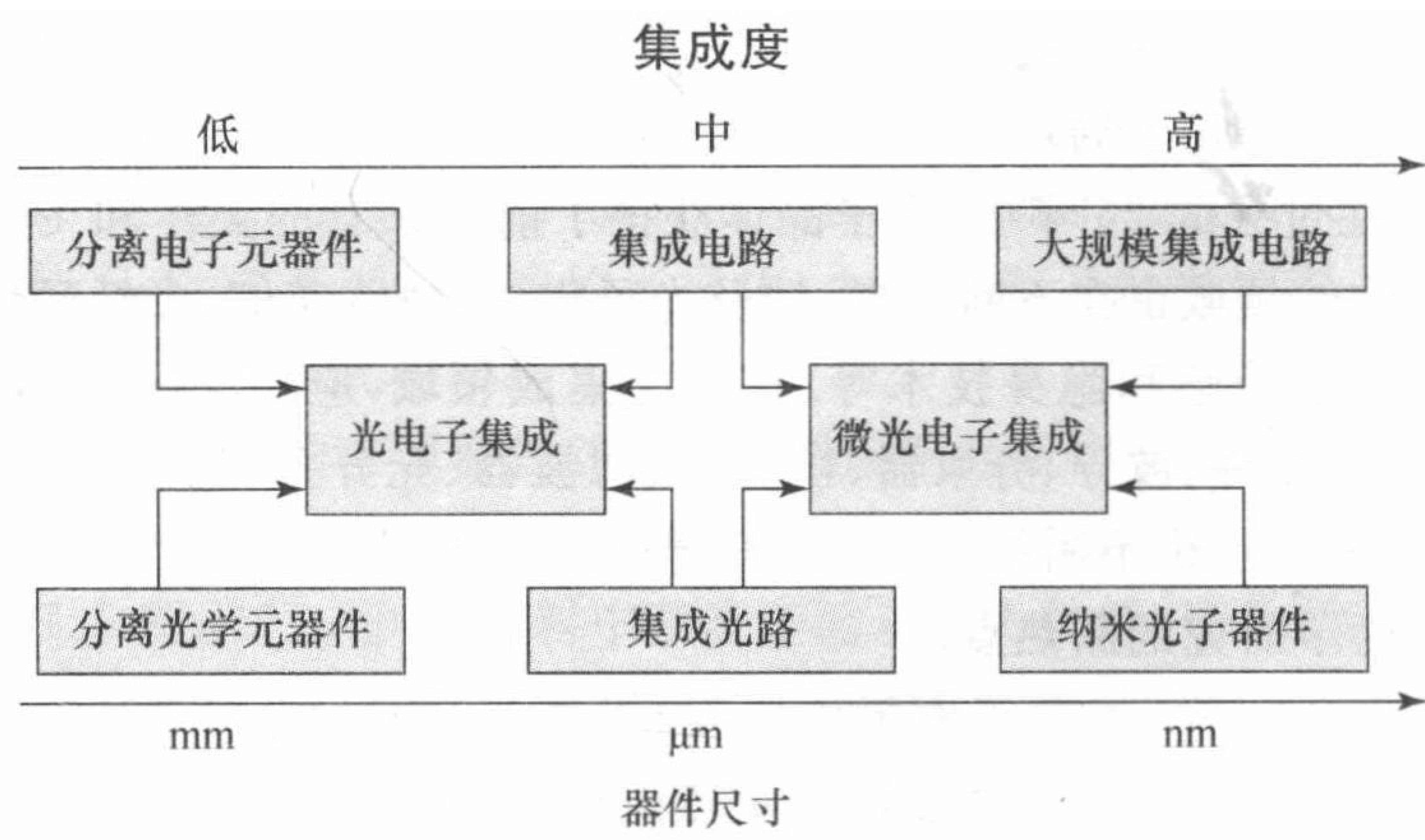


图 1.6 微电子学与集成光学相结合并形成微纳光电子学的必然趋势

人们预测，如能真正实现光子器件和微电子器件相结合的硅基光电子集成，同时利用光子和电子作为信息和能量的载体，就可以实现超大容量、超高速率、低热耗散和低串扰的信息传输，而这将会对未来的通信和计算机技术产生革命性的影响[4~8].

众所周知，硅基电子集成工艺已经颇为成熟. 而与之相比，硅基光电子集成工艺却有着相当大的难度，其主要限制体现在以下三个方面：一是光子器件和电子器件的结构复杂，两者在结构设计上存在着能否相互兼容的问题；二是制作工艺繁杂，存在着各种工艺和前后工序之间能否相互兼容的问题；三是电互连、光互连与光耦合等问题. 因为在光电子集成中，要实现光的互连，必须通过各种波导器件来实现，这自然就又出现了如何集成波导结构的问题. 为了能实现高集成度、高性能和实用化的硅基光电子集成电路，需要统筹考虑和解决材料选择、结构设计、制作工艺和性能检测等一系列问题. 结构设计与制作工艺的相容性问题则是能否实现硅基光子集成的关键所在. 科学家们已提出了两种可供参考的集成方案，即光电混合集成和单芯片集成. Luxtera 公司在该领域率先取得了突破性进展. 2006 年 8 月 22 日，Luxtera 推出业界第一款采用标准 CMOS 工艺的单芯片集成光电器件. Luxtera 的突破性技术将光学芯片和主流的电子技术集成到一起，实现光纤同电子芯片的直接配合，采用标准的 0.13 μm 绝缘体上硅(SOI)CMOS 工艺，进而还可以实现数字逻辑芯片同光芯片的集成，大大降低光器件的尺寸、功耗和成本. Luxtera 的这一成就让我们看到了实现全硅光电集成的曙光.

图 1.7 为 Intel 公司提出的单片硅基光电子集成所需的几个单元器件[32]. 这些器件覆盖了光耦合、光传输、滤波等无源器件及光发射、光调制、光探测等有源器件领域，具有了对光信息进行各种处理的功能.

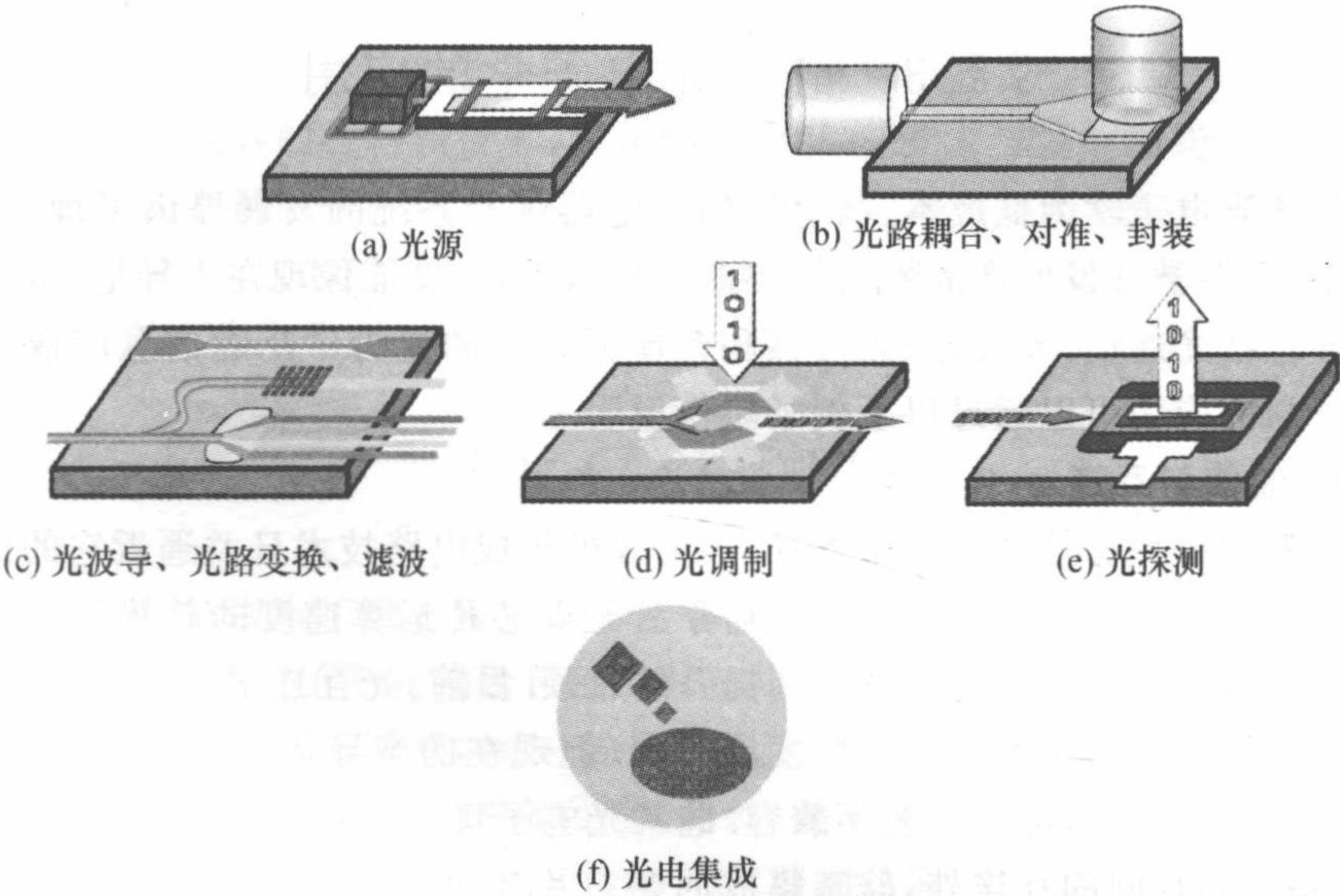


图 1.7 Intel 硅基集成光路几大组成模块

然而，其核心挑战就是如何将这些单元器件“硅片化”并集成起来，即利用硅或与硅兼容的其他材料，应用硅工艺，在同一硅基底上同时制作若干微纳量级、以光子、电子为载体的信息功能器件，形成一个完整的具有综合功能的新型芯片单元. 一旦这一目标得以实现，未来的硅基光电子集成芯片将如图 1.8 所示，从而使高密度数据通信、实时传感、高速控制等现代信息社会的需求在小型化、低成本、高可靠的基础上获得满足.

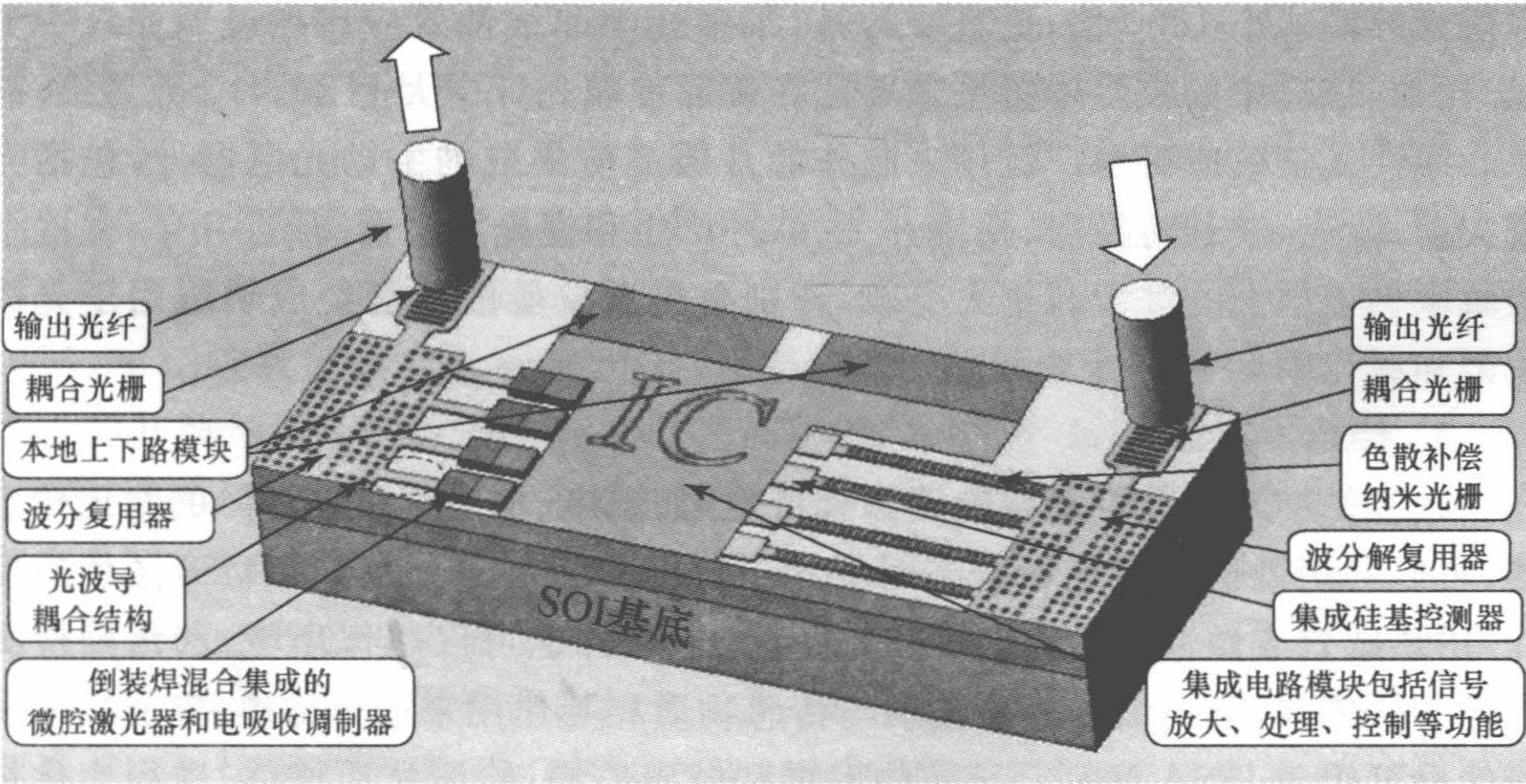


图 1.8 未来硅基光电子集成芯片示意图

§1.3 硅基光电子学的应用

硅基光电子学为低成本、高容量的光电器件及系统的发展提供了理论和实验基础.如果被开发的产品性能达到商业化水平,硅就能像现在主导电子产业一样主导包括光通信、光互连、光传感和光能源在内的更为广泛的信息产业.硅基光电子学的应用可以通过以下例子略作说明.

1. 通信与互连

随着微处理器性能呈指数增长,超大规模集成电路技术日益逼近它的极限,计算机系统内部通信速度和宽带落后于处理器芯片运算速度的趋势日益扩大,铜互连将成为计算机系统整体性能提升的瓶颈.目前,光互连被公认是实现计算机间乃至芯片间高速通信的重要发展方向.但现在的半导体发光器件多用化合物材料制备,与硅微电子工艺不兼容.硅基光电子集成正是解决光互连的最佳途径.除基于芯片间的互连外,硅基集成回路芯片还可以实现板与板之间的互连、机栈与机栈之间的互连、远程主干网上光纤通信系统的互连等.对于通信领域来说,硅基光电子首先在远距离传输中有很重要应用,分立器件的替换能够节省数十亿成本;其次,更多的有着更高集成度的器件与系统将在短距离通信领域得到应用.

2. 非线性光学效应

现在在硅基光电子学的研究当中,一个重要的分支就是研究各种非线性光学效应来制造有源光学器件.由于硅和硅氧化物的折射率有很大差别,光可以比玻璃光纤或氧化硅波导更紧的限制在硅波导中.在硅基波导中光学模式尺寸比单模光纤模式小 1000 倍;更重要的是,很多光学非线性效应比在玻璃光纤中更强.比如,在硅中的拉曼增益系数要比在普通玻璃光纤中大上 3～4 个数量级.硅基非线性光学效应可以广泛用来制作芯片级高度集成的主动光电器件,包括光放大器、激光、波长转换器、超快开关、脉冲产生和慢光发生器.最近几年,在这个领域取得了很大的进步和重大突破,特别是受激拉曼散射已经成功地用于光放大器和激光中.

3. 传感

硅基微纳米器件可用做检测物理特性(如温度、加速度、压力等)的物理传感器,可以利用微机电系统(micro-electro-mechanical systems, MEMS)做加速度计,用热敏效应做温度传感器.同时硅基光子器件也可以用作环境、水污染检测的化学传感器,利用波导倏逝波场与传感物质相互作用来进行传感.集成光电子器件也可用于 DNA 序列乃至药物发展和健康监控,检测抗原抗体,进行生化传感检测等.集成硅基光电子具有把不同的传感单元集成在一个芯片上增加芯片

的探测处理能力等优势.

4. 片上实验室

利用与硅兼容的材料和技术,将化学或生化传感器与其他的光学、电学、力学元器件集成在同一个芯片上,集探测、传感、滤波、频谱分析、数据处理于一体,以获得便携式、低成本的片上实验室. 这种片上实验室在疾病诊断、DNA 分析、免疫检测、环境监控等方面具有极大的应用潜力.

§1.4 总 结

信息时代的微电子技术按照摩尔定律飞速发展,但随着特征尺寸减小到十几个纳米以下,微电子产业能否再依照摩尔定律前进已面临挑战. 此时,半导体最小单元——晶体管的栅极(gate)的尺寸将达到 5 nm,电子通过量子隧穿效应(tunneling effect),将会自行穿越信道. 此外,随着 IC 集成度的进一步提高,器件的功耗和散热也成为了制约微电子技术发展的难题. 另一方面,基于计算机与通信网络化的信息技术,希望微电子器件具有更快的处理速度、更大的数据存储容量和更高的传输速率. 然而,传统金属(铜)互连技术由于 RC 延迟和热耗散造成互连"瓶颈",基于电子作为信息载体的硅集成电路技术将难以满足以上的要求. 为此,将微电子和光电子结合起来,充分发挥硅基微电子先进成熟的工艺技术、高密度集成及价格低廉以及光子极高带宽、超快传输速率及高抗干扰性的优势,已经成为了信息技术发展的必然和业界的普遍共识. 光子技术和微电子技术相结合,通过光子布线技术,使得微电子的数据传输速度和数据传输量极大增加,集成电路的功耗得到有效的降低.

目前硅基光电子学和光子集成技术已经发展到了一个崭新的阶段,新技术新器件层出不穷,有关硅的各种传统观念被新的实验结果一一突破. 2004 年 Intel 公司在 *Nature* 上报道了调制带宽超过 1 GHz 的硅基高速光波导调制器,同年康奈尔大学的研究人员在《自然》上发表了基于 SOI 波导微环谐振器的亚纳秒级全光开关. 2005 年 Intel 公司又研制出了 1550 nm 光泵浦的激射波长为 1686 nm 的连续硅基 Raman 激光器,康奈尔大学的研究人员在《自然》上报道了基于 SOI 波导微环谐振器的高速电光调制器. 2006 年 Intel 公司和加州大学联合研制成功了世界上首个电泵浦连续激射硅基Ⅲ-Ⅴ族混合集成激光器,丹麦的研究人员在《自然》上报道了应变硅的强电光效应. 2007 年 Intel 公司将硅基电光调制器的 3 dB 带宽扩展到 30 GHz,实现了 40 Gbit/s 的信号传输,达到目前商用Ⅲ-Ⅴ族和铌酸锂外调制器的水平. 2008 年美国 Luxtera 公司向世人展示了世界上第一块在 130 nm CMOS 生产线上制造的硅基单片集成高速 CMOS 光子收发模块,采用波分复用(wavelength division multiplexing, WDM)技术,数据

传输速率 4×10 Gbit/s；同年 Intel 公司报道了采用调制器阵列和波分复用 MUX/DeMUX 的硅基光电子集成芯片，数据传输速率为 200 Gbit/s.

从基于硅材料的微电子工业本身来考虑，众多的专业公司使得硅基光电集成更加具备了坚实的工业背景和光明的市场前景. 因此，发展硅基光电子集成，将微电子和光电子技术结合起来，在集成电路芯片间乃至芯片内部引入集成光路，是信息工业发展的必由之路.

参考文献

[1] Intel Corporation. Moor's law, made real by Intel innovations. http://www.intel.com/technology/mooreslaw/index.htm.

[2] ITRS. International technology roadmap for semiconductors 2009 edition. http://www.itrs.net/Links/2009ITRS/Home2009.htm.

[3] 王启明. 支撑光网络发展的光子集成器件研发与趋势. 深圳大学学报：理工版，2004，21(3)：189.

[4] 侯金，郜定山，周治平. 光子晶体器件进展. 光学技术，2009，35(1)：93.

[5] 周治平，郜定山，汪毅，等. 硅基集成光电子器件的新进展. 激光与光电子学进展，2007，44(2)：31.

[6] 周治平，王兴军，冯俊波，等. 硅基微纳光电子系统中光源的研究现状及发展趋势. 激光与光电子学进展，2009，46(10)：28.

[7] 郜定山. 二维光子晶体的理论及器件研究[博士后出站报告]. 武汉：华中科技大学，2007.

[8] Miller S E. A survey of integrated optics. IEEE J. Quant. Electr., 1972, QE-8: 199.

[9] Soref R A. Silicon-based optoelectronics. Proc. IEEE, 1993, 81: 1687.

[10] Paniccia M. Silicon photonics applications research results & integration challenges. Berkley Lecture Series: 2004-04-27.

[11] 王阳元，黄如，刘晓彦，等. 面向产业需求的 21 世纪微电子技术的发展. 物理，2004，33(6—7).

[12] Soref R A, Bennett B R. Electrooptical effects in silicon. IEEE J. Quant. Electr., 1987, QE-23: 123.

[13] Liu A S, Jones R, Liao L, et al. A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor. Nature, 2004, 427: 615.

[14] Mao A, Liu J, Gao D, Zhou Z. Silicon phase modulator based on n-p-n configuration. Electr. Lett., 2008, 44: 438.

[15] Manipatruni S, Xu Q F, Lipson M. PINIP based high-speed high-extinction ratio micron-size silicon electro-optic modulator. Opt. Exp., 2007, 15: 13035.

[16] Rong H, Liu A, Jones R, et al. An all-silicon Raman laser. Nature, 2005, 7023: 292.

[17] Isshiki H, de Dood M J A, Kimura T. Self-assembled infrared-luminescent Er-Si-O crystallites on silicon. Appl. Phys. Lett., 2004, 85: 4343.

[18] Wang X J, Kimura T, Zhou Z. Enhanced Er^{3+} luminescence of Ersilicate by Y and Yb codoping. Proc. 6th IEEE International Conference on Group IV Photonics, 2009.

[19] Wang X J, Nakajima T, Kimura T. Fabrication and characterization of Er silicates on SiO_2/Si substrates. Appl. Phys. Lett., 2009, 95(4): 1906.

[20] Liu J F, Sun X C, Pan D, et al. Tensile-strained, n-type Ge as a gain medium for monolithic laser integration on Si. Opt. Exp., 2007, 15(18): 11272.

[21] Kasap S O. Optoelectronics and photocics principles and practices. New jersey: Prentice-Hall International, 2001.

[22] Kang Y M, Liu H D, Morse M, et al. Monolithic germanium/silicon avalanche photodiodes with 340GHz gain-bandwidth product. Nature Photonics 2008, 44(4): 378.

[23] Yablonovitch E. Inhibited spontaneous emission in solid-state physics and electronics. Phys. Rev. Lett., 1987, 58: 2059.

[24] John S. Strong localization of photons in certain disordered dielectric superlattices. Phys. Rev. Lett., 1987, 58: 2486.

[25] Bozhevolnyi S I, Volkov V S, Devaux E, et al. Channel plasmonsubwavelength metal components including interferomers and ring resonators. Nature, 2006, 440: 508.

[26] Ebbesen T W, Genet C, Bozhevolnyi S I. Surface plasmon circuitry. Phys. Today, 2008, 61: 44.

[27] Koller D M, Hohenau A, Ditlbacher H, et al. Organnicligth-emitting diodes as surface plasmon emitters. Nature Photonics, 2008, 2: 684.

[28] Walters R J, van Loon R V A, Brunets I, et al. A silicon-based electrical source of surface plasmonpolaritons. Nature Materials, 2010, 9: 21.

[29] Wang Y, Zhou Z. Strong enhancement of erbium ion emission by a metallic double-grating. Appl. Phys. Lett., 2006, 89(25): 18.

[30] Wang Y, Zhou Z. Silicon optical amplifier based on surface-plasmon-polariton enhancement. Appl. Phys. Lett., 2007, 91(05): 3504.

[31] 周治平. 硅基微米/纳米级光电子器件及集成技术的基础研究. 给中国科学技术协会的973建议, 2004, 11.

[32] Intel Corporation. Silicon Photonic Technology. http://www.intel.com/go/sp.

第二章　硅中光子与电子的相互作用

电子学是研究控制电子流的学科，而光子学是研究控制光子流的学科. 光子和电子在半导体中相互作用：光子激发可移动的电子，电子产生并控制光子流. 光子与电子的相互作用构成了光电子学的基础. 研究硅中光子与电子相互作用的学科因此被定义为硅基光电子学. 半导体材料被用来制作光电子器件，如光源（发光二极管和激光）、放大器、调制器、探测器、传感器和非线性光学器件等，已有相当长的历史. 近年来，半导体光电子器件与硅基电子器件在尺寸、能耗和工艺兼容性方面的突破推动了硅基光电子学的快速发展.

§2.1　波动光学与光子光学[1]

波动光学属于经典光学的范畴，经典电磁理论能够对其作出完备的解释. 然而，经典电磁理论是有缺陷的，有些光学现象根本无法解释. 经典理论的缺陷促进了人们对电磁理论的进一步深化，进而诞生了量子电磁理论，即量子电动力学. 在量子电动力学框架下，电场 $\boldsymbol{E}$ 和磁场 $\boldsymbol{H}$ 被数学处理为向（矢）量空间中的算符，并假定它们满足特定算符方程和对易关系，因此确定了时变性质和相互关系. 量子电动力学更为准确地描述了电磁场与物质的相互作用. 利用量子电动力学推导出的结果在本质上具有量子性，因而无法用经典理论解释.

利用量子电动力学获得的能够描述粒子性、空间位置和电磁场能量微扰的简单关系式，我们可以推导出很多光的量子特性以及光与物质的作用原理. 这就是光子光学. 光子光学能够解释经典理论所无法解释的光学现象. 但是，光子光学也并不能解释所有的光学现象.

本节将介绍光子的概念及其特性，阐述控制光子能量、动量、偏振、位置、时间和干涉的一系列规律并讨论光子流的性质.

2.1.1　光子及其特性

光由粒子组成，这些粒子称为光子. 单光子静止质量为零，但具有电磁能量和动量. 它还具有决定偏振特性的本征角动量（或称自旋）. 光子在真空中以光速（c_0）传播，在物质中会降低传播速度. 光子同时具有波动特征，这些特征决定了光子的位置性质及其干涉和衍射规律.

光子的概念来源于普朗克为解决长期悬而未决的黑体辐射谱问题所做的理

论尝试. 通过量化辐射黑体腔中各电磁模式所允许的能量值，普朗克成功地解决了这一难题. 为了阐述光子的概念以及光子光学的基本规律，我们首先考虑光在谐振腔(如空腔)中的现象. 这一假定将光的传播空间简化为几何形状，但并未对问题的实质有任何显著影响，由此推出的结论同样适用于没有谐振腔的情形.

根据电磁理论，光在体积为 V 且无损耗的谐振腔中的状态，可由不同频率、偏振和空间分布的离散正交的电磁模式之和表示：

$$\boldsymbol{E}(\boldsymbol{r},t)=\sum_q A_q E_q(\boldsymbol{r})\exp(\mathrm{i}2\pi\nu_q t)\hat{\boldsymbol{e}}_q,\quad q=1,2,3,\cdots,\tag{2.1.1}$$

其中 q 阶电磁模式具有复振幅 A_q，频率 ν_q，平行于单位矢量 $\hat{\boldsymbol{e}}_q$ 的偏振方向. 其空间分布由复函数 $E_q(\boldsymbol{r})$表示，并通过$\int_q |E_q(\boldsymbol{r})|^2\mathrm{d}\boldsymbol{r}=1$ 进行归一. 展开项 $E_q(\boldsymbol{r})$ 和单位矢量 $\hat{\boldsymbol{e}}_q$ 的选取并非唯一. 在立方形特征尺寸为 d 的谐振腔中，往往选取驻波集合作为空间展开函数族，如下所示：

$$E_q(\boldsymbol{r})=\left(\frac{2}{d}\right)^{3/2}\sin\frac{q_x\pi x}{d}\sin\frac{q_y\pi y}{d}\sin\frac{q_z\pi z}{d},\tag{2.1.2}$$

其中 q_x,q_y 和 q_z 是整数，它们一起被索引矢量表示为 $\boldsymbol{q}=(q_x,q_y,q_z)$[详见图 2.1(a)]. 该模式包含的能量为

$$E_q=\frac{1}{2}\epsilon\int_V \boldsymbol{E}(\boldsymbol{r},t)\cdot\boldsymbol{E}(\boldsymbol{r},t)\mathrm{d}\boldsymbol{r}=\frac{1}{2}\epsilon|A_q|^2.\tag{2.1.3}$$

在经典电磁理论中，能量 E_q 可为任意非负值，不论有多小. 总能量是各模式的能量之和.

如上所述的电磁理论在光子光学中依然成立，只是对每个电磁模式的能量大小加以限制. 如图 2.1(b)所示能量不再为任意值，模式能量的可能值为等间距的非负数组. 因此，模式能量被量化，成为某固定值的整数倍. 这一固定的能量单位则由光子携带.

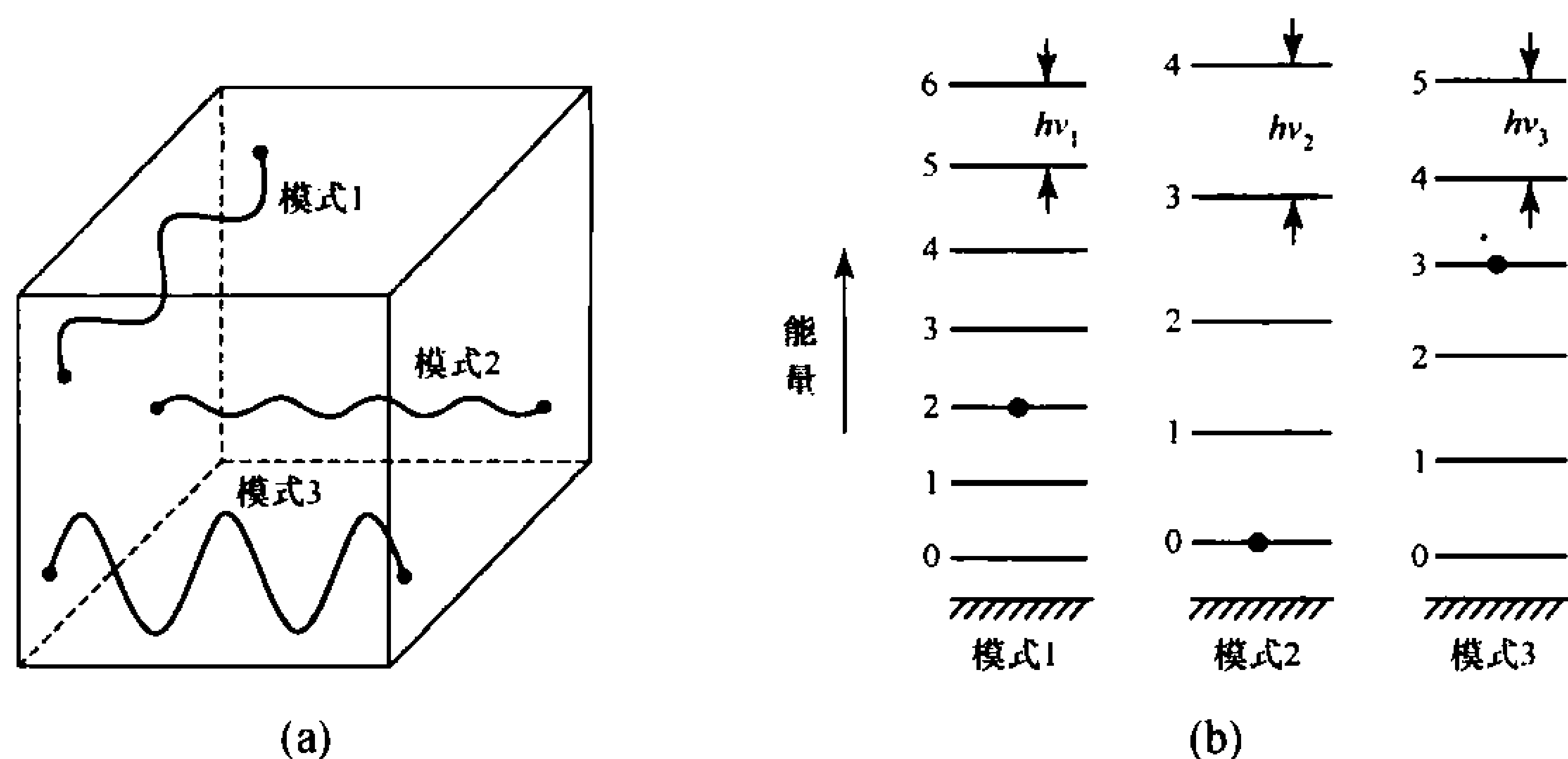


图 2.1　立方谐振腔内频率和方向不同的模式(a)；三个模式允许的频率(和对应能量)ν_1,ν_2,ν_3. 黑点表示每个模式内的光子数；模式 1,2,3 分别包含 2,0,3 个光子(b)

如图 2.2 所示波长为 1 cm 的光子倒波长为 1 cm^{-1}. 光子频率 $\nu=3\times10^{14}$ Hz，对应波长 $\lambda_0=1\ \mu m$，能量 1.24 eV 以及倒波长 10 000 cm^{-1}.

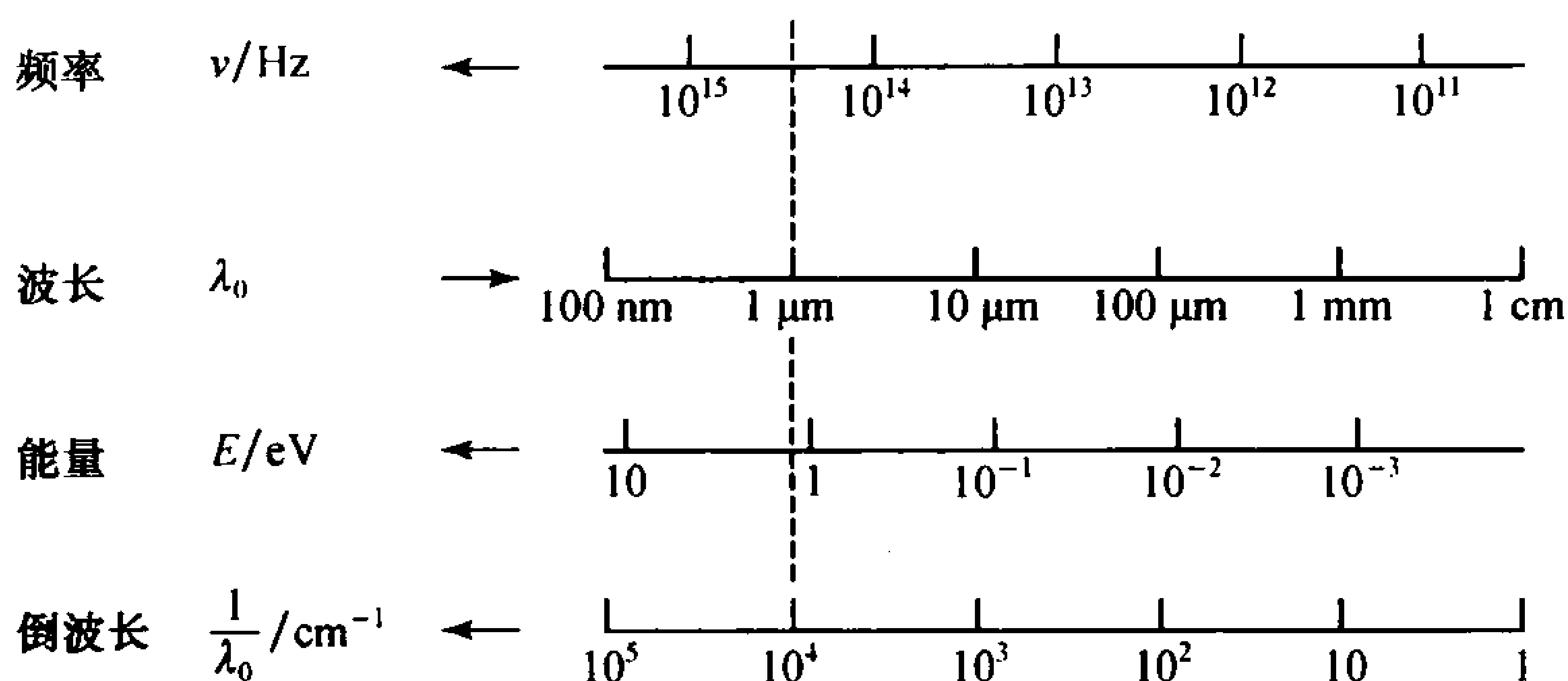


图 2.2 光子频率 ν，波长 λ_0，能量 E 和倒波长 $1/\lambda_0$ 间的关系

综上所述，光子光学认为光在谐振腔中由一系列模式构成，每个模式都包含整数个相同光子. 不同模式的特征，如频率、空间分布、传播方向和偏振方向，成为其光子的属性.

2.1.2 光子的偏振

如前所述，光可以表示为不同频率、方向和偏振的基本模式之和. 光子的偏振是其所代表的模式的偏振. 但是，模式集合的选取并不唯一. 以下是从光子光学的角度对光的偏振性质的解释.

1. 线性偏振光子

考虑一束光由两个沿 z 方向传递的平面波模式叠加而成，一个模式沿 x 方向线性偏振，另一个模式沿 y 方向线性偏振：

$$\boldsymbol{E}(\boldsymbol{r},t) = (A_x\hat{\boldsymbol{x}} + A_y\hat{\boldsymbol{y}})\exp(-\mathrm{i}kz)\exp(\mathrm{i}2\pi\nu t). \tag{2.1.4}$$

同样，同一电磁场也可以在另一坐标系 $Ox'y'$ 中表示(如将原先坐标系 Oxy 旋转 45°得到的新坐标系)，如图 2.3 所示. 因此，我们可以将原先的电磁场看做是沿 x' 和 y' 方向偏振的两个模式之和，即

$$\boldsymbol{E}(\boldsymbol{r},t) = (A_{x'}\hat{\boldsymbol{x}}' + A_{y'}\hat{\boldsymbol{y}}')\exp(-\mathrm{i}kz)\exp(\mathrm{i}2\pi\nu t), \tag{2.1.5}$$

其中

$$A_{x'} = \frac{1}{\sqrt{2}}(A_x - A_y), \quad A_{y'} = \frac{1}{\sqrt{2}}(A_x + A_y).$$

假设一个光子占据了沿 x 方向偏振的模式，而沿 y 方向偏振的模式没有光子，那么有多少概率发现一个沿着 x' 方向偏振的光子？在光子光学中，我们通过常规概率论方法加以解决. 观测到沿 x,y,x',y' 偏振光子的概率分别正比于 $|A_x|^2$，$|A_y|^2$，$|A_{x'}|^2$，$|A_{y'}|^2$. 在本例中，$|A_x|^2=1$，$|A_y|^2=0$

所以$|A_{x'}|^2=|A_{y'}|^2=1/2$.因此,如果有一个光子沿 x 方向偏振而没有光子沿 y 方向,则发现沿 x',y'方向偏振光子的概率均为 1/2,如图 2.4 所示.

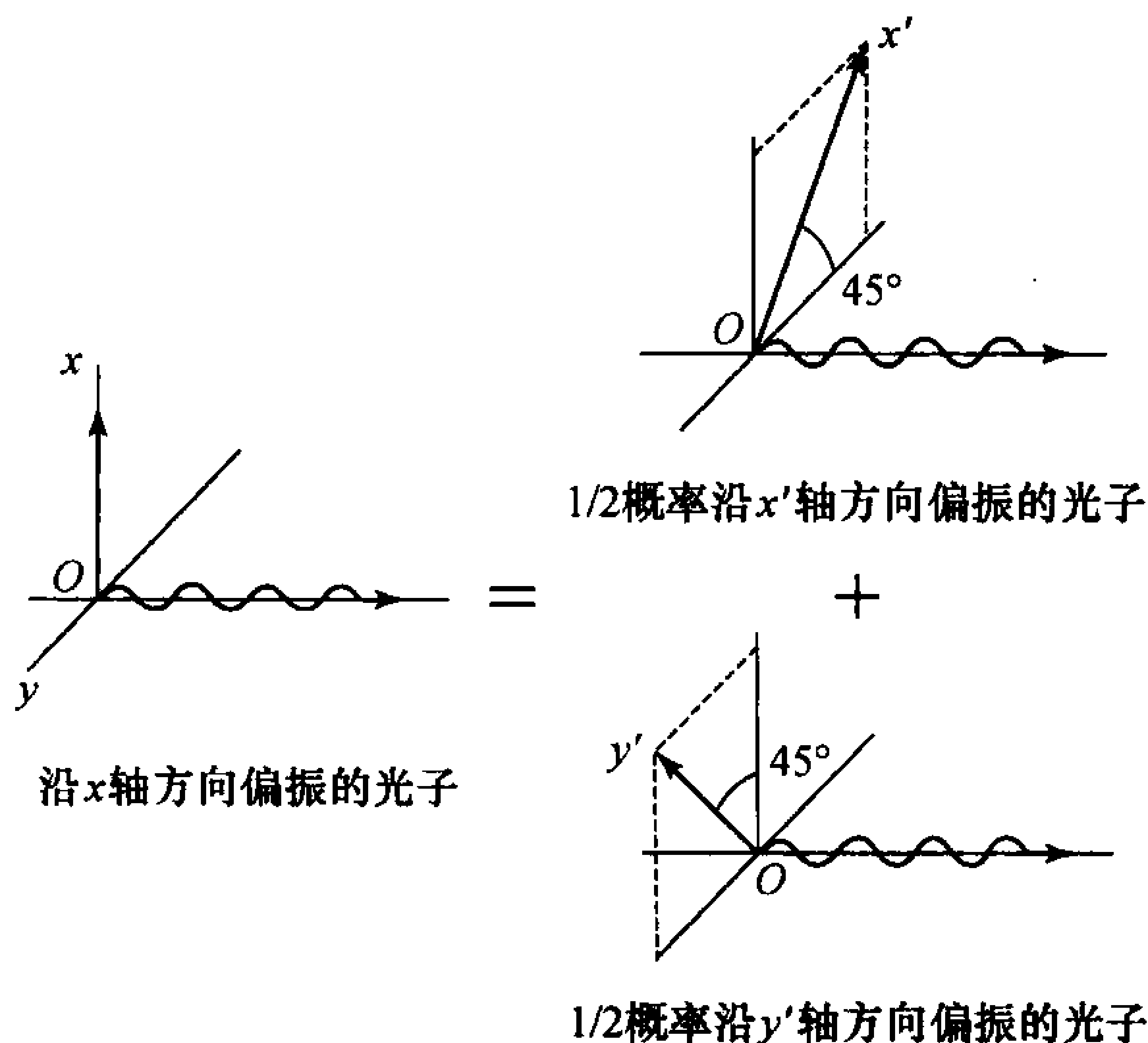


图 2.3 将线性偏振光依概率分解为两束光

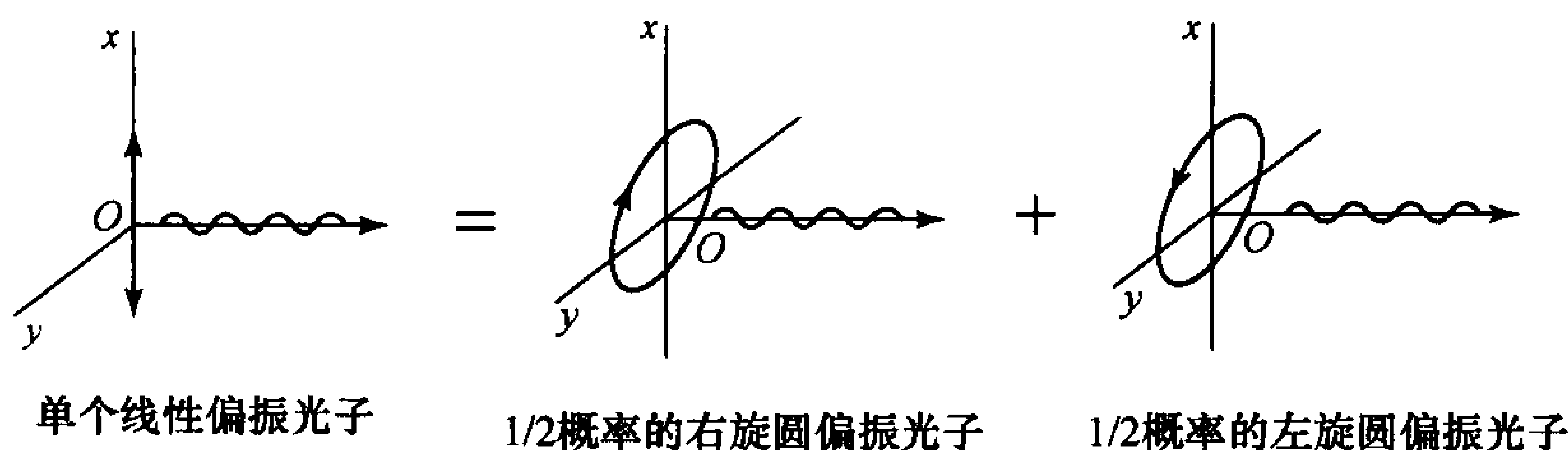


图 2.4 单线偏振光子等价于 1/2 概率的右旋圆偏振光子和 1/2 概率的左旋圆偏振光子之和

2. 圆偏振光子

基于两个圆偏振平面波模式(右旋圆偏振和左旋圆偏振)的展开式可以表达为

$$\boldsymbol{E}(\boldsymbol{r},t) = (A_{\mathrm{R}}\hat{\boldsymbol{e}}_{\mathrm{R}} + A_{\mathrm{L}}\hat{\boldsymbol{e}}_{\mathrm{L}})\exp(-\mathrm{i}kz)\exp(\mathrm{i}2\pi\nu t), \tag{2.1.6}$$

其中 $\hat{\boldsymbol{e}}_{\mathrm{R}}=(1/\sqrt{2})(\hat{\boldsymbol{x}}+\mathrm{i}\hat{\boldsymbol{y}})$,$\hat{\boldsymbol{e}}_{\mathrm{L}}=(1/\sqrt{2})(\hat{\boldsymbol{x}}-\mathrm{i}\hat{\boldsymbol{y}})$.这些模式分别携有右旋和左旋圆偏振光子.同样,发现左、右旋圆偏振光子的概率也正比于振幅$|A_{\mathrm{R}}|^2$ 和$|A_{\mathrm{L}}|^2$.正如图 2.4 所示,一个线性偏振光子等价于左旋和右旋圆偏振光子以各 1/2 概率的叠加.相反,当一个圆偏振光子通过线偏光镜时,能够通过的概率为 1/2.

3. 光子自旋

光子存在本征角动量(自旋).光子自旋的大小被量化为两个值:

$$S = \pm \hbar. \tag{2.1.7}$$

右旋(或左旋)圆偏振光子的自旋矢量平行(或反平行)于其动量矢量. 线性偏振光子等概率的表现为平行自旋和反平行自旋. 光子能够传递给物体线性动量,同样,圆偏振光子也能在物体上产生力矩. 比如,一个圆偏振光子可以在半波长的石英片上产生作用力矩.

2.1.3 光子动量

光子动量与自身波函数的波矢相关,并遵循如下规律: 某一模式的光子由平面波表示为

$$\boldsymbol{E}(\boldsymbol{r},t) = A\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\exp(\mathrm{i}2\pi\nu t)\hat{\boldsymbol{e}}, \tag{2.1.8}$$

具有动能矢量

$$\boldsymbol{p} = \hbar\boldsymbol{k}. \tag{2.1.9}$$

光子依照波矢方向传播,其动量大小为

$$p = \hbar k = \hbar 2\pi/\lambda, \tag{2.1.10}$$

即

$$p = \frac{h}{\lambda}. \tag{2.1.11}$$

对于平面波,波动光学可推出相同的能量-动量关系 $\boldsymbol{p}=(E/c)\hat{\boldsymbol{k}}$,其中 $\boldsymbol{p}$ 是每单位体积波的动量,E 是单位体积的能量,$\hat{\boldsymbol{k}}$ 是方向平行于 $\boldsymbol{k}$ 的单位矢量. 但是,在波动光学中不存在"光子"这一概念,因此带有 $\hbar$ 的公式(2.1.9)和(2.1.10)只适用于光子光学.

比平面波更具有一般性的局部波是以复波函数 $A\boldsymbol{E}(\boldsymbol{r})\exp(\mathrm{i}2\pi\nu t)\hat{\boldsymbol{e}}$ 来表达的波. 它可以通过傅里叶变换展开为具有不同波矢的平面波之和. 含有波矢为 $\boldsymbol{k}$ 的项可写为 $A(\boldsymbol{k})\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\exp(\mathrm{i}2\pi\nu t)\hat{\boldsymbol{e}}$,其中 $A(\boldsymbol{k})$为幅度.

由任意复波函数 $A\boldsymbol{E}(\boldsymbol{r})\exp(\mathrm{i}2\pi\nu t)\hat{\boldsymbol{e}}$ 描述的光子动量是不确定的. 它以正比于$|A(\boldsymbol{k})|^2$ 的概率具有动量

$$\boldsymbol{p} = \hbar\boldsymbol{k}, \tag{2.1.12}$$

其中 $A(\boldsymbol{k})$是以平面波傅里叶展开 $E(\boldsymbol{r})$后含有波矢 $\boldsymbol{k}$ 的那一项的幅度.

平面波单光子具有确定动量(确定方向和大小),但其位置是完全不确定的; 在 $z=0$ 平面内任意位置可等概率观测到该光子. 当平面波光子通过小孔时,它的位置被限定在小孔内,而动量方向变得不确定. 与平面波光子相对应的是球面波光子. 它的位置完全固定(在球波的球心处),而动量方向完全不确定.

光子的动量是守恒的,这就意味着发射单光子的原子将会受到大小为 $h\nu/c$ 的后坐力. 此外,光子动量可以传递给具有有限质量的物体,产生力和导致机械运动. 比如,光束可以使垂直于自身传播方向的原子束发生偏转,我们称这种现

象为辐射压力.

2.1.4　光子能量和位置

光子光学认为,电磁模式能量的量化应参照以单光子能量为间距的等差数列.单光子在频率为 ν 的模式中的能量,即光子能量,可以表述为

$$E = h\nu = \hbar\omega, \tag{2.1.13}$$

其中 $h=6.63\times10^{-34}$ J·s 是普朗克常量,$\hbar=h/2\pi$.这一模式的能量增减变化只能以 $h\nu$ 为单位.

然而,如果某个模式不携有光子,其能量为 $E_0=\frac{1}{2}h\nu$,称为零点能.当该模式携有 n 个光子时,模式能量为

$$E_n = \left(n+\frac{1}{2}\right)h\nu, \quad n = 0,1,2,\cdots. \tag{2.1.14}$$

因为光子频率和它所携有的能量成正比,随着辐射频率的增加,光的粒子性逐渐突显.此外,随着波长变短,光的干涉和衍射这一类波动特性也变得难以辨识.X 射线和 Γ 射线总表现得像粒子束,而不像波状的无线电波.光学领域中的光频率使得光既有粒子性也有波动性,因此带动光子光学的发展.

光子位置与光的强度密切相关.每个光子都与一个波相联系,该波由电磁模式的复函数 $A\boldsymbol{E}(\boldsymbol{r})\exp(\mathrm{i}2\pi\nu t)\hat{\boldsymbol{e}}$ 所描述.然而,当一个光子撞向位置为 $\boldsymbol{r}$ 且垂直于传播方向的面元为 $\mathrm{d}\boldsymbol{A}$ 的探测器时,光子的不可分割性使得它要么被探测到要么被忽略.光子被探测到的位置并不确定,它受光强 $I(\boldsymbol{r})\propto|\boldsymbol{E}(\boldsymbol{r})|^2$ 影响,同时遵从如下的光子位置定律:

无论何时,在位置 $\boldsymbol{r}$ 的面元 $\mathrm{d}\boldsymbol{A}$ 处观测到光子的概率 $p(\boldsymbol{r})\mathrm{d}\boldsymbol{A}$ 正比于在该处的光强 $I(\boldsymbol{r})\propto|\boldsymbol{E}(\boldsymbol{r})|^2$,即

$$p(\boldsymbol{r})\mathrm{d}\boldsymbol{A} \propto I(\boldsymbol{r})\mathrm{d}\boldsymbol{A}. \tag{2.1.15}$$

光子在光强较强的地方被探测到的概率较大.比如,某一电磁模式的光子是强度分布 $I(x,y,z)\propto\sin^2(\pi z/d)$ 的驻波,其中 $0\leqslant z\leqslant d$.那么,$z=d/2$ 处最容易探测到光子,而在 $z=0$ 或 $z=d$ 处不可能探测到光子.光子不是在空间中无线延展的波,也不是具有固定位置的粒子,而是表现为在固定位置处有限空间内活动的实体,这一表现被称为波粒二象性.光子能够被探测到是光子粒子性的显著表现.

光子的这种统计规律可通过单光子透过分光镜的传输现象来进一步描述.理想的分光镜是将入射光无损的均分为两束互成直角的出射光.其特征参量包括透射系数 T 和反射系数 $R=1-T$.透射光光强 I_t 和反射光光强 I_r 可以利用入射光光强 I 和电磁关系 $I_r=(1-T)I, I_t=TI$ 计算得到.

单个光子是不可分割的，所以它必须在分光镜允许的两个传播方向内选其一. 入射在分光镜上的光子选择结果遵从光子位置随机规律(2.1.15). 光子能够透射的概率正比于 I_t，因此正比于透射系数 T，光子能够反射的概率也因此正比于 $1-T$. 从概率论的角度说，这一问题等同于掷硬币得到的正反面，图 2.5 解释了这一过程.

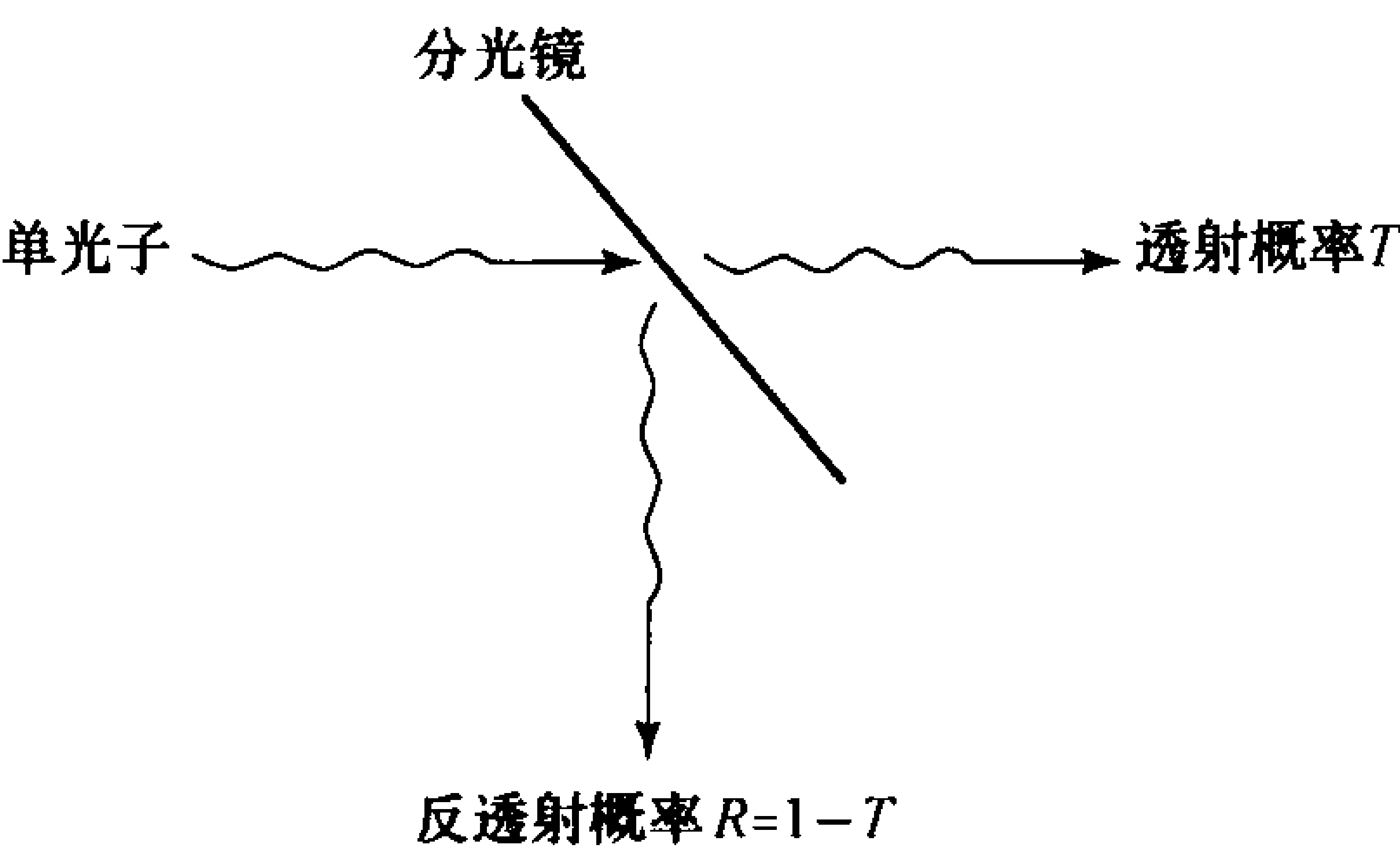


图 2.5 光子在分光镜处的随机透射或反射

2.1.5 光子的干涉

杨氏双孔干涉实验可证明光的波动性. 然而，即便让光子一个一个地通过干涉仪，我们依然可以重复出杨氏双孔实验的结果. 这一结果可利用光子光学中光子位置定律加以解释. 利用波动光学我们能够计算出最后观察面上的强度分布，这一分布被转化为光子的随机密度函数，表示其可能出现的位置. 光子的干涉同样可以归结于两条路径的相位差.

假设一平面波照向带有两个小孔的平板，如图 2.6 所示. 在板后形成两个球面波，它们在观察面上相互干涉. 根据菲涅耳近似，所产生的强度分布呈正弦波状

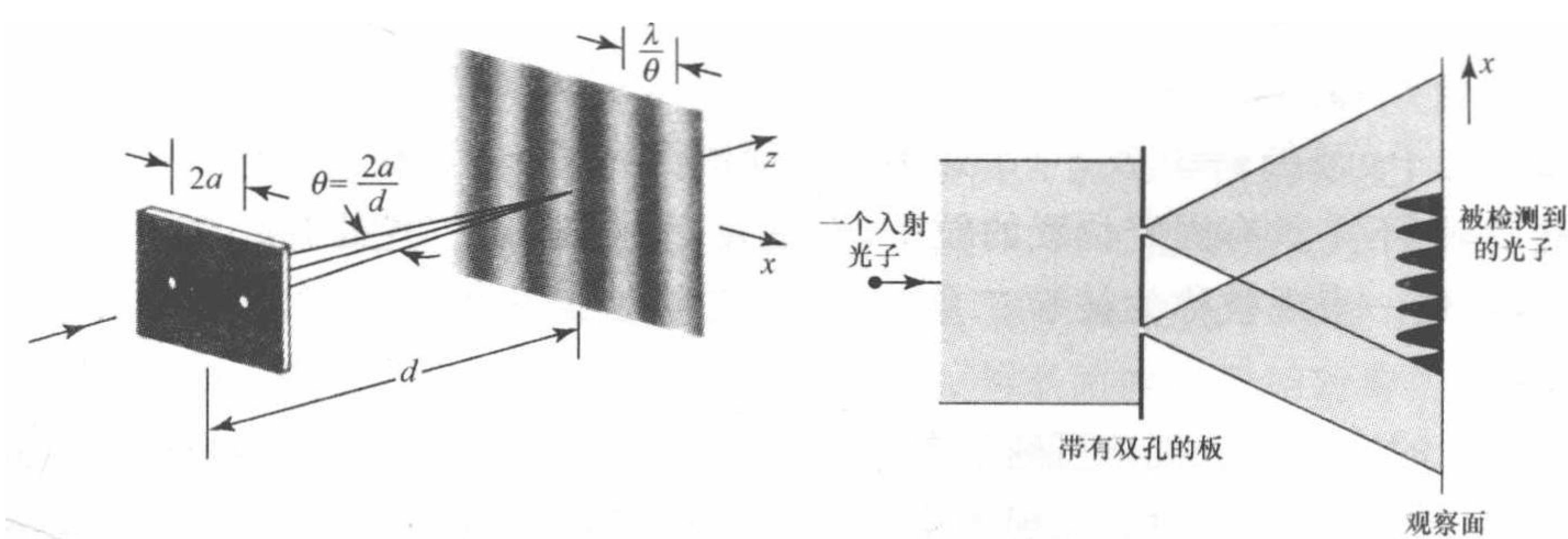


图 2.6 单光子的杨氏双孔实验，干涉图样 $I(x)$ 正比于在 x 处探测到光子的概率

$$I(x)=2I_0\left(1+\cos\frac{2\pi\theta x}{\lambda}\right),\tag{2.1.16}$$

其中 I_0 是两束球面波到达观察面时的强度，λ 是波长，θ 为两孔到观察面某点所成的角度．两孔的连线是 x 轴．当入射光足够强时，式(2.1.16)描述了观察面上的强度分布．

现在如果在干涉仪中只有一个光子，在位置 x 处观测到光子的概率正比于 $I(x)$，如式(2.1.15)所述．在 $I(x)$ 最大值处最容易观测到光子，在 $I(x)=0$ 处则永远不会．正如泰勒在 1909 年进行的实验，重复单光子双孔实验并将观测到光子的位置绘制在直方图上，所得结果依然为经典的杨氏双孔干涉图像．在光子光学中，干涉分布代表光子在观察面上被探测到的概率分布．

干涉结果的产生来自于光子位置的不确定性，使得单光子能够同时通过干涉仪的两孔．因而，单光子在观察面上的概率分布能够反映实验干涉本质．如果干涉仪两孔中堵住一孔，干涉图像将会消失，因为光子只能从另一孔通过，无法反映出干涉本质．

2.1.6　光子的时域特性

光子的时间位置关系：式(2.1.1)给出了单色(单频)模式的表达式，该表达式在时域是无限延展的．单色模式中的光子在任何时候都能等概率地被探测．然而，如前所述，谐振腔内(外)辐射模式表达式并不唯一，利用多色模式(如有限时域的波包)可以得到更为普遍的模场展开式．在任意位置时间间隔 t 和 $t+\mathrm{d}t$ 之间，能由波函数 $\boldsymbol{E}(\boldsymbol{r},t)$ 描述的光子，其被探测到的概率正比于 $I(\boldsymbol{r},t)\mathrm{d}t\propto|\boldsymbol{E}(\boldsymbol{r},t)|^2\mathrm{d}t$.

光子位置定律[参见式(2.1.14)]因此可以包括光子时间局域效应：在点 $\boldsymbol{r}$ 处面元 $\mathrm{d}\boldsymbol{A}$ 上，在时间间隔 t 和 $t+\mathrm{d}t$ 间观察到光子的概率，正比于在 $\boldsymbol{r}$ 处和 t 时的模场强度，即

$$p(\boldsymbol{r},t)\mathrm{d}\boldsymbol{A}\mathrm{d}t\propto I(\boldsymbol{r},t)\mathrm{d}\boldsymbol{A}\mathrm{d}t\propto|U(\boldsymbol{r},t)|^2\mathrm{d}\boldsymbol{A}\mathrm{d}t.\tag{2.1.17}$$

光子的时间能量不确定性：能够探测到频率为 ν 单色模式光子的时机是完全不确定的，然而光子频率 ν(和光子能量 $h\nu$)却是确定的．另一方面，处于强度为 $I(t)$ 宽度为 σ_t 的波包模式中的光子，一定会被限制在 σ_t 这一时间段内．依据傅里叶变换性质，伴随对光子时域的限制，光子频率(和光子能量)变得不确定．其结果是该光子成为多色光子．频率的不确定性可由 $\boldsymbol{E}(t)$ 傅里叶展开中的谐波分量决定：

$$\boldsymbol{E}(t)=\int_{-\infty}^{\infty}\boldsymbol{V}(\nu)\exp(\mathrm{i}2\pi\nu t)\mathrm{d}\nu,\tag{2.1.18}$$

其中 $\boldsymbol{V}(\nu)$ 是 $\boldsymbol{E}(t)$ 的傅里叶变化．这里出于简化，略去了光子在各位置 $\boldsymbol{r}$ 处

的各向异性. $|V(\nu)|^2$ 的展宽 σ_ν 代表谱宽. 如果 σ_t 是函数 $|\boldsymbol{E}(t)|^2$ 的均方根值(如能量的均方根值),那么 σ_t 和 σ_ν 满足时域带宽反比关系,即 $\sigma_\nu\sigma_t \geqslant 1/(4\pi)$,或 $\sigma_\omega\sigma_t \geqslant 1/2$.

光子能量 $h\nu$ 的最小分辨率为 $\sigma_E = h\sigma_\nu$. 同理,光子能量的不确定度和它可能被探测到的时段,必须满足

$$\sigma_E\sigma_t \geqslant \frac{\hbar}{2}, \tag{2.1.19}$$

这就是时间能量不确定性关系. 它类似于光子位置和波数(动量)间的关系,在同时测定光子位置和动量的精度上加以限制. 对于多色光子,平均能量 $\bar{E} = h\bar{\nu} = \hbar\bar{\omega}$.

总而言之,单色光子($\sigma_\nu \to 0$)在任何时刻($\sigma_t \to \infty$)都可能被观测到,时域不受限. 相反,光学波包中的光子在时域上受限,因此频率不确定,频谱展宽,导致能量的不确定. 因此,波包光子可以看做运动着的在空间有所展宽的能量包.

2.1.7 小结

电磁辐射可以表示为各模式之和,如均一单色平面波之和

$$\boldsymbol{E}(\boldsymbol{r},t) = \sum_q A_q \exp(-\mathrm{i}\boldsymbol{k}_q \cdot \boldsymbol{r})\exp(\mathrm{i}2\pi\nu_q t)\hat{\boldsymbol{e}}_q. \tag{2.1.20}$$

每个平面波都有两种正交偏振态(如垂直/水平线偏振态,左手/右手圆偏振态等),由 $\hat{\boldsymbol{e}}_q$ 表示. 当测量模式能量大小时,能量值是整数倍个能量子(光子). 模式 q 中的光子具有如下性质:

① 能量 $E = h\nu_q$;

② 动量 $\boldsymbol{p} = \hbar\boldsymbol{k}$;

③ 如果是圆偏振光,自旋 $S = \pm\hbar$;

④ 如果模式波函数是单色平面波,那么无论何时何地观测到光子的概率是相同的.

模式的选取不唯一. 可用非单色光(准单色光)、非平面波进行展开,展开式为

$$\boldsymbol{E}(\boldsymbol{r},t) = \sum_q A_q E_q(\boldsymbol{r},t)\hat{\boldsymbol{e}}_q, \tag{2.1.21}$$

此时,模式 q 中的光子具有如下性质:

① 光子位置和时间局域效应由波函数 $E_q(\boldsymbol{r},t)$ 决定. 在点 $\boldsymbol{r}$ 处面元 $\mathrm{d}\boldsymbol{A}$ 上,时间间隔 t 到 $t+\mathrm{d}t$ 内,观测到光子的概率正比于 $|E_q(\boldsymbol{r},t)|^2\mathrm{d}\boldsymbol{A}\mathrm{d}t$.

② 如果 $E_q(\boldsymbol{r},t)$ 在时域以 σ_t 有限展宽,即如果光子时域有限,那么光子能量 $h\nu_q$ 具有不确定关系 $h\nu_q \geqslant h/(4\pi\sigma_t)$.

③ 如果 $E_q(\boldsymbol{r},t)$ 在横截面($z=0$)上空间位置有限展宽,即如果光子在 x 轴

上有固定位置(y 轴依然不确定),那么光子动量方向不确定.光子动量的展宽可由 $E_q(\boldsymbol{r},t)$ 的平面波展开式决定,其中含有波矢 $\boldsymbol{k}$ 的项对应着光子动量 $\hbar\boldsymbol{k}$.在横截面上光子空间位置不确定性减小使得光子动量方向上不确定增大.

§2.2　半导体能带结构[1]

半导体具有能带和带隙,并在允许的能级间,遵循原子中产生光子的原理,吸收和发射光子.在本节中,我们将简述对半导体光子学中较为重要的半导体的性质,特别是它的能带结构及其变化规律.半导体是导电性介于金属和绝缘体之间的晶体或非晶体固体,其导电性能随着温度、掺杂浓度以及光照射强度的变化而显著变化.半导体材料独特的能级结构使得其具有特殊的电学和光学特性,而这些特性在研究光子与电子的相互作用时尤为重要.

2.2.1　能带和载流子

1. 能带

固体材料中的原子具有很强的相互作用,因此它们不能被简单地看成是独立个体的集合.价带电子并不属于独立的原子,相反,它们应属于由原子组成的整个系统.在由原子构成的晶格结构所形成的周期性势场中,薛定谔方程关于电子能量的解导致分离的原子能级,并形成半导体的能带结构.每个能带中有大量的离散能级,我们能近似将其认为是准连续的.价带和导带被宽为 E_g(参见图 2.7)的"禁止"能量间隙所分离.E_g 被称为带隙能量,它在决定半导体材料的电学和光学性能中有着举足轻重的作用.价带填满、宽带隙(>3 eV)的材料被称为电绝缘体;带隙较小或不存在的材料被称为电导体.半导体材料的带隙大概在 0.1~3 eV 间.

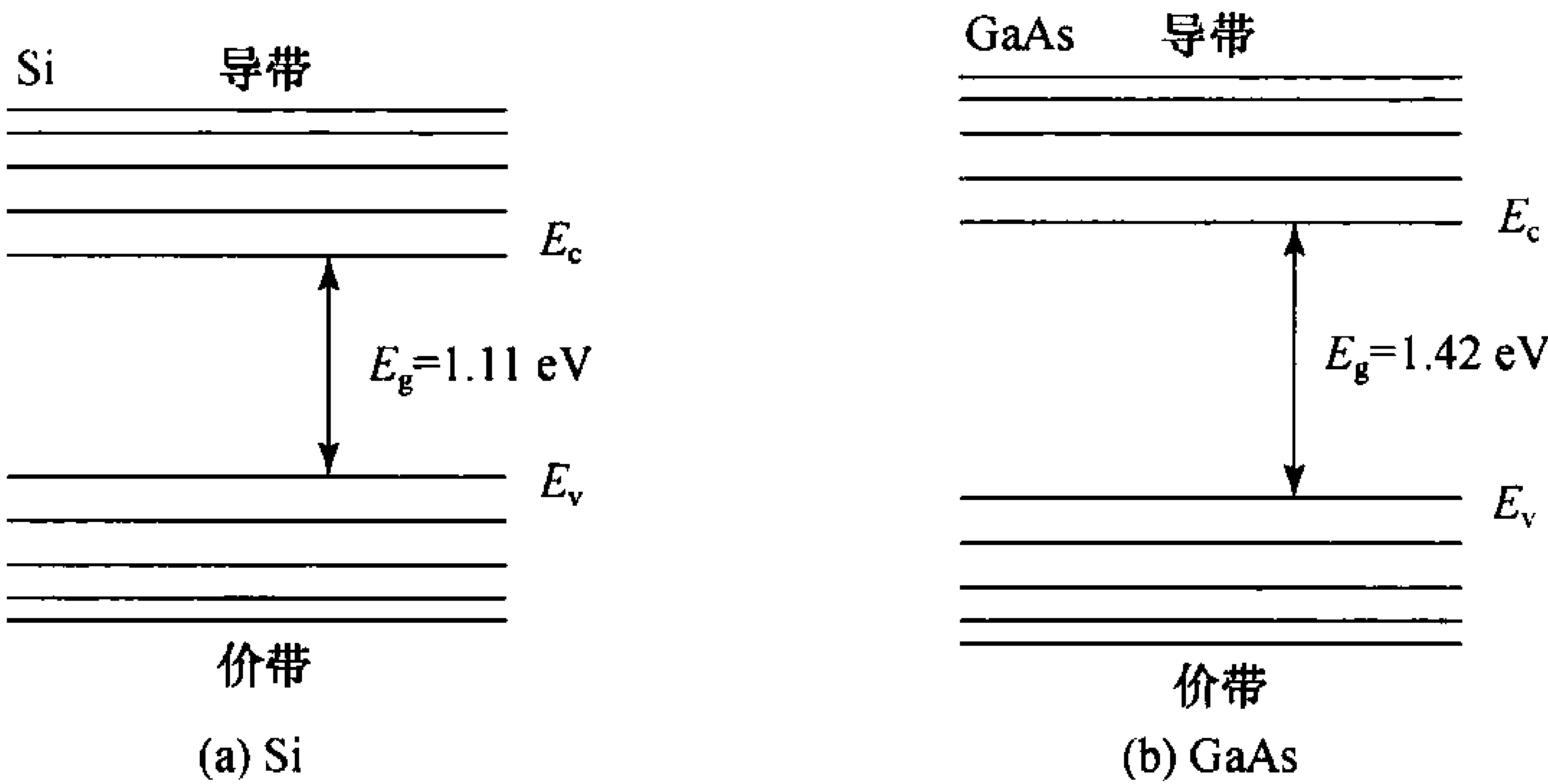


图 2.7　Si 和 GaAs 的能带

2. 载流子

根据泡利不相容原理,两个电子不能占据相同的量子态,低的能级先被电子填充. 在单元素半导体中,例如 Si 和 Ge,每个原子有四个价电子;价带具有一定数量的量子态使得在没有热激发的情况下价带正好被完全填充,而导带完全空置. 在这种情况下,材料是不能导电的. 当温度上升时,一些电子被热激发到有充足空置量子态的导带中(参见图 2.8). 在导带中,电子成为自由载流子;在外电场的作用下它们可以在晶格中漂移,为产生电流作贡献. 另一方面,从价带跃至导带的电子在价带留下一个空的量子态,使得留在价带的电子在外场的作用下可以改变位置,从而留在价带的电子集合在整体上产生迁移. 这个过程也可被等效的看成是离去电子留下空穴的反方向运动. 空穴可被认为是一个带正电荷 $+e$ 的载流子. 每个电子激发过程都将在导带中产生自由电子和在价带中产生自由空穴. 两种载流子可以在外场作用下漂移从而产生电流. 半导体材料在热激发下载流子数量上升,电导率显著增加.

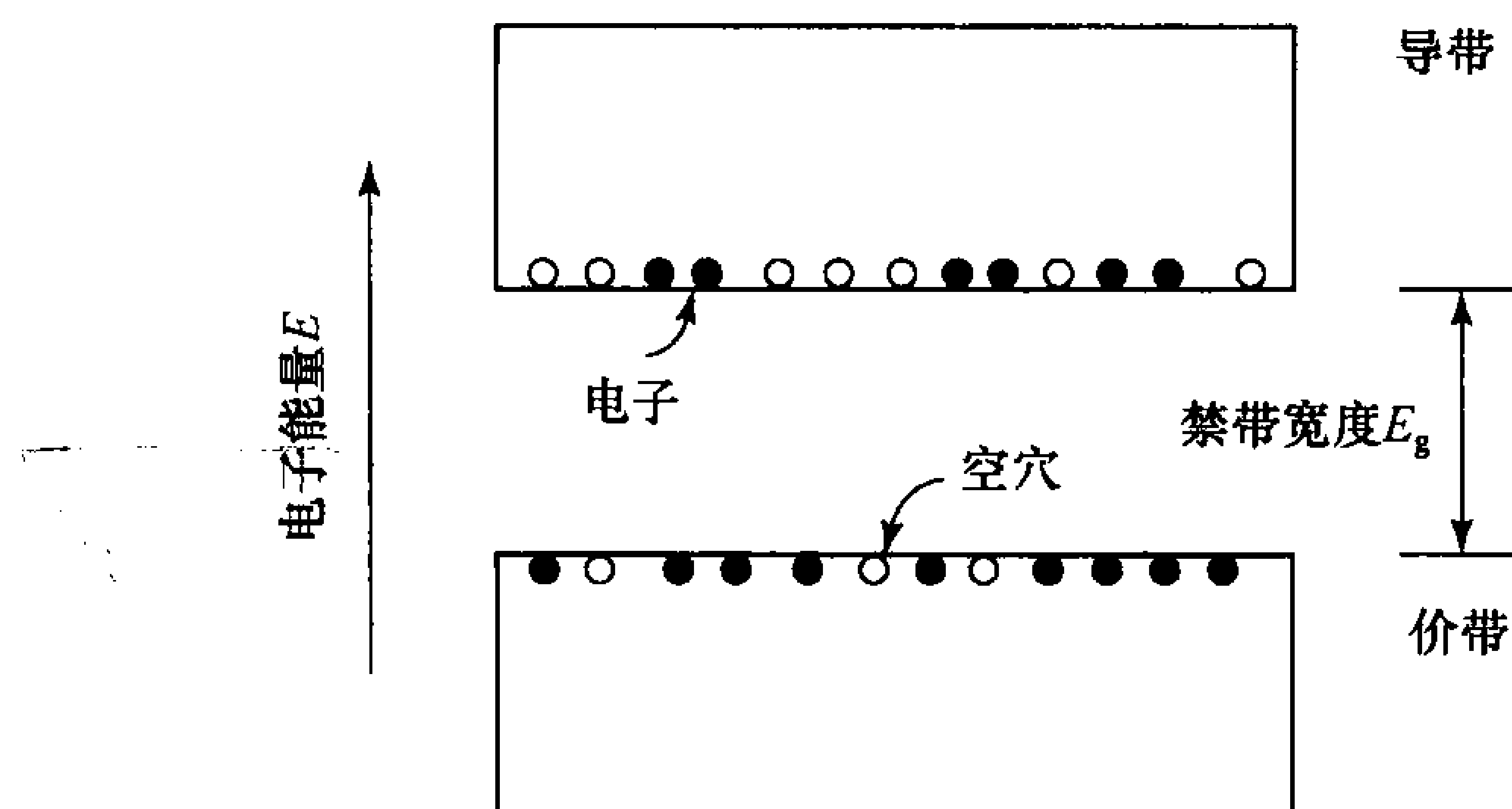


图 2.8 $T>0$ K 时导带中的电子和价带中的空穴

3. 能量-动量关系

自由空间中电子的能量 E 和动量 $\boldsymbol{P}$ 有如下关系:

$$E = P^2/(2m_0) = \hbar^2 k^2/(2m_0),$$

其中 P 是动量的模,k 是传播矢量 $\boldsymbol{k}$ 的模 ,根据波动方程有 $\boldsymbol{k}=\boldsymbol{P}/\hbar$,$m_0$ 是电子的静止质量($m_0=9.1\times10^{-31}$ kg). E-k 关系是一个简单的抛物线.

半导体导带中的电子和价带中的空穴的运动受不同情况的制约. 它们都遵循薛定谔方程和材料周期性晶格结构的限制. E-k 关系是传播矢量 $\boldsymbol{k}$ 的三个分量(k_1,k_2,k_3)周期函数,周期为$\left(\frac{\pi}{a_1},\frac{\pi}{a_2},\frac{\pi}{a_3}\right)$,其中 a_1,a_2,a_3 是晶体的晶格常数. 图 2.9 是 Si 和 GaAs 的 E-k 关系沿 $\boldsymbol{k}$ 的两个不同方向的截面图. 导带电子的能量不仅仅由其动量的模 k 决定,也与其在晶体中的运动方向有关.

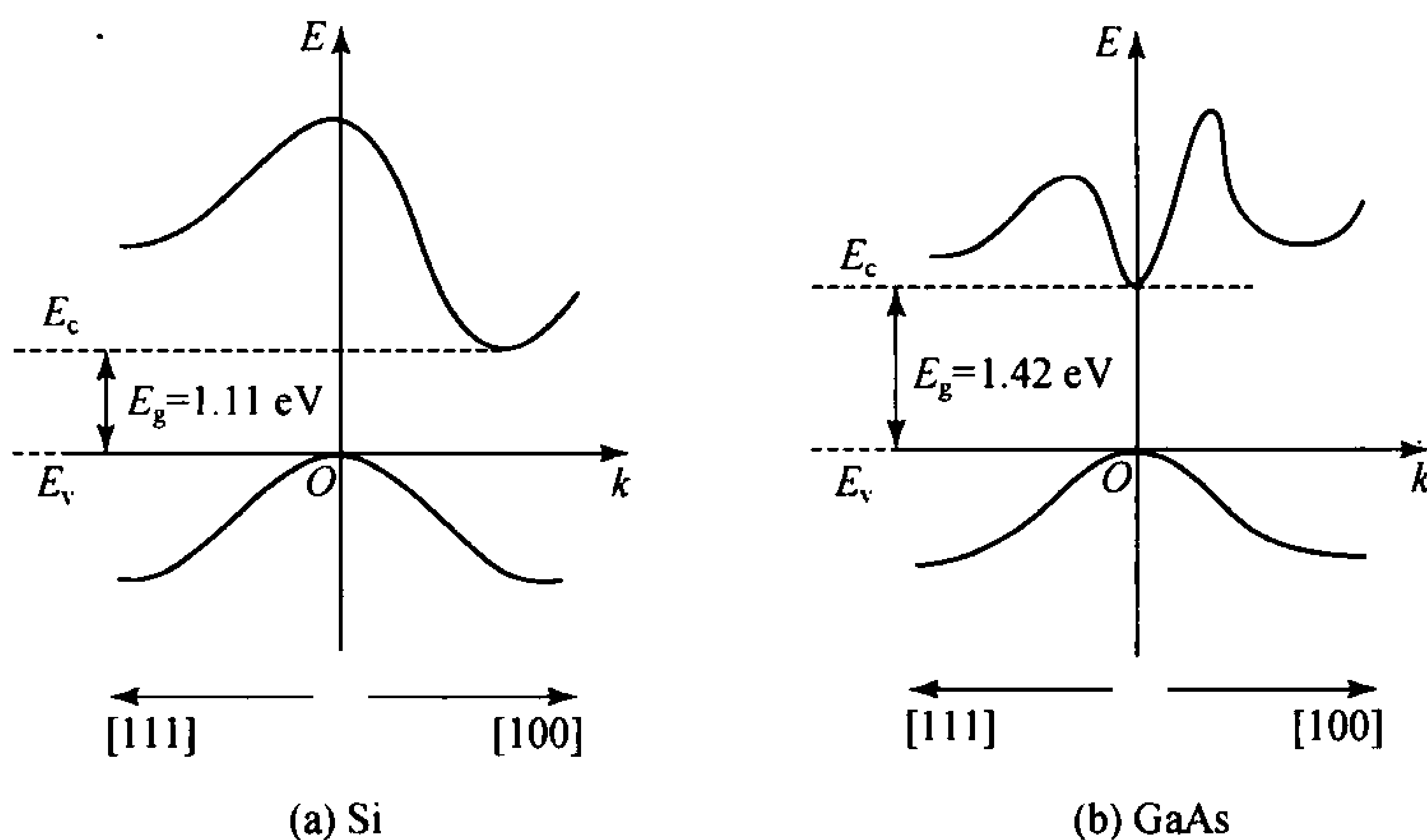


图 2.9　在 Si 和 GaAs 中沿[111]和[100]晶体方向的 E-k 关系截面图

在导带底 E-k 关系可以近似被认为是抛物曲线

$$E = E_c + \frac{\hbar^2 k^2}{2m_c}, \tag{2.2.1}$$

其中 E_c 是导带底的能量，m_c 是一个常数，被称为电子在导带中有效质量(参见图 2.10).

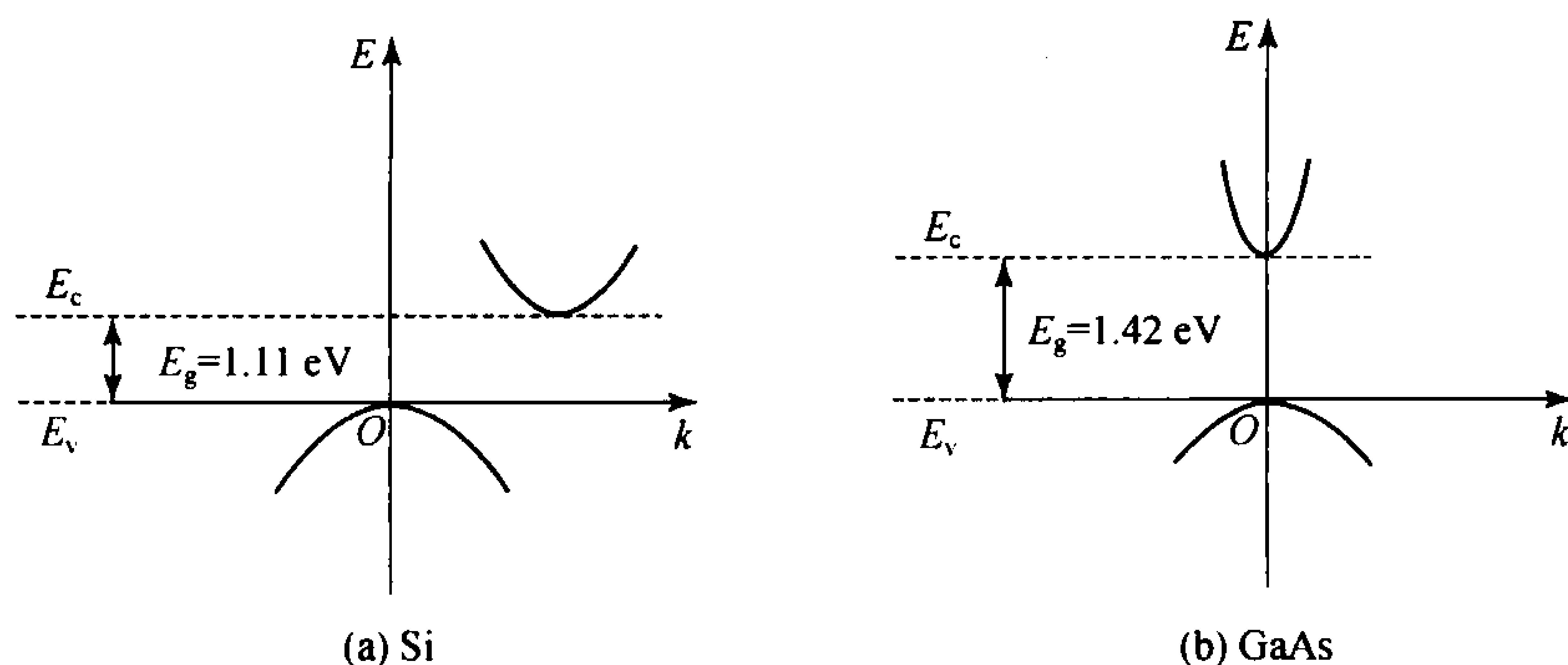


图 2.10　Si 和 GaAs 在导带底和价带顶近似是抛物曲线的 E-k 关系

类似地，在价带顶

$$E = E_v - \frac{\hbar^2 k^2}{2m_c}, \tag{2.2.2}$$

其中 E_v 是价带顶的能量，m_v 是空穴在价带中的有效质量. 总之，有效质量由晶向和能带决定. Si 和 GaAs 的平均有效质量和电子静止质量 m_0 比值在表 2.1 中列举出来.

表 2.1 Si 和 GaAs 中电子和空穴的平均有效质量

	m_v/m_0	m_v/m_0
Si	0.33	0.5
GaAs	0.07	0.5

4. 直接带隙和间接带隙半导体

价带顶和导带底能量对应相同的动量(k 相同)的半导体称为直接带隙材料,反之则称为间接带隙材料.这两种材料的重要区别是,电子在间接带隙材料中价带顶和导带底的转换过程中需要显著的动量变化.从图 2.10 中,我们看到 Si 是间接带隙材料,GaAs 是直接带隙材料.我们将在后面的篇幅里阐述为什么直接带隙的材料发光效率较高,而间接带隙的发光效率则相对较低.

半导体的电学和光学特性可以通过添加少量杂质(或称为掺杂)来改变,从而形成掺杂半导体.载流子浓度在掺杂后会在数量级上发生变化.导致多余价电子的掺杂被称为施主杂质,代替原来晶体中少量的原子,自由电子成为材料中主要的载流子,这样的材料被称为 n 型半导体.Ⅴ族元素(P 或 As)代替Ⅳ族单元素(Si 或 Ge)半导体会生成 n 型半导体.类似地,p 型半导体材料是由于加入了使得价电子不足的掺杂(受主杂质)而形成,自由空穴成为材料中主要的载流子.Ⅲ族元素(B 或 In)代替Ⅳ族单元素半导体会生成 p 型半导体.

没有被掺杂的半导体称为本征半导体材料,掺杂的半导体称为非本征半导体材料.自由电子和空穴在本征半导体中浓度相同,$n=p=n_i$,其中 n_i 随着温度上升指数级增加.n 型半导体中的自由电子浓度(多子)远远高于自由空穴的浓度(少子),即 $n \gg p$.反之则在 p 型半导体中成立,空穴是多子,即 $p \gg n$.在室温下掺杂的半导体的多子浓度基本等于掺杂浓度.

2.2.2 载流子浓度

要想得到载流子(电子和空穴)浓度与能量的函数关系,首先需要了解每个能级允许的状态密度(态密度)和每个能级被占据的概率.

1. 态密度

半导体材料中电子的量子态由其能量 E、传播矢量 $\boldsymbol{k}$[其模 k 和 E 大致遵循式(2.2.1)或(2.2.2)]和电子自旋所表征,并由满足一定边界条件的波函数所描述.

在导带带边的电子能近似地被描述为,质量为 m_c 的粒子被限制在边界完全反射的三维立方体(量纲 d)中,例如理论上可以处理为三维无限深势阱.薛定谔方程在该三维立方体中的驻波解需要满足如下条件:传播矢量 $\boldsymbol{k}=[k_x,k_y,k_z]$有着分立的取值 $\boldsymbol{k}=[q_1\pi/d,\ q_2\pi/d,\ q_3\pi/d]$,其中 q_1,q_2,q_3 为正整数.传播矢量 $\boldsymbol{k}$ 必

须在晶格点上，其立方单元晶胞的边长是 π/d. 因此在 $\boldsymbol{k}$ 空间中单位体积内有 $(d/\pi)^3$ 格点. 传播矢量 $\boldsymbol{k}$ 的模量在$[0,k]$间的量子态数应该是在半径为 k 球体内的正八分之一体中的格点数，正八分之一体的体积大概为 $(1/8)4\pi k^3/3=\pi k^3/6$. 由于电子自旋量子数有两个可能取值，在 $\boldsymbol{k}$ 空间的每点都对应着两个量子态. 在 d^3 空间中约有

$$\frac{2\left(\frac{\pi k^3}{6}\right)}{\left(\frac{\pi}{d}\right)^3}=\left(\frac{k^3}{3\pi^2}\right)d^3,$$

即在单位体积中有 $(k^3/3\pi^2)$ 个格点. 在波数 $k\sim k+\Delta k$ 间的量子态数是

$$\rho(k)\Delta k=\left[\frac{\mathrm{d}}{\mathrm{d}k}\left(\frac{k^3}{3\pi^2}\right)\right]\Delta k=\left(\frac{k^2}{\pi^2}\right)\Delta k,$$

因此态密度为

$$\rho(k)=\left(\frac{k^2}{\pi^2}\right). \tag{2.2.3}$$

电磁波传播模式中，场的偏振有两个自由度（光子的两个自旋值），同样，在半导体中电子也存在着两个由自旋决定的状态. 在谐振光学中，允许的传播矢量 $\boldsymbol{k}$ 的解被表示为线性的频率-波数关系 $\nu=\frac{ck}{2\pi}$. 在半导体物理学中，允许的传播矢量 $\boldsymbol{k}$ 的解依据式(2.2.1)和(2.2.2)被表示为允许的能量-波数关系.

如果 $\rho_c(E)\Delta E$ 代表导带中单位体积能量在 E 至 $E+\Delta E$ 的能级数，根据式(2.2.1)中 E 和 k 的一对一关系，态密度应有 $\rho_c(E)\mathrm{d}E=\rho(k)\mathrm{d}k$. 因此导带中单位能量的态密度函数为

$$\rho_c(E)=\frac{\rho(k)}{\frac{\mathrm{d}E}{\mathrm{d}k}}.$$

同样，价带中能量态密度应满足

$$\rho_v(E)=\frac{\rho(k)}{\frac{\mathrm{d}E}{\mathrm{d}k}},$$

其中 E 和 k 的关系满足式(2.2.2). 利用式(2.2.1)和(2.2.2)在导带和价带边的二次近似 E-k 关系，可以推导出价带和导带中的 $\mathrm{d}E/\mathrm{d}k$ 关系，结果如下

$$\rho_c(E)=\frac{(2m_c)^{3/2}}{2\pi^2\hbar^3}(E-E_c)^{1/2},\quad E\geqslant E_c, \tag{2.2.4}$$

$$\rho_v(E)=\frac{(2m_v)^{3/2}}{2\pi^2\hbar^3}(E_v-E)^{1/2},\quad E\leqslant E_v, \tag{2.2.5}$$

1/2 次方关系是由于电子和空穴在带边的能量-波数的 2 次方关系决定的. 图 2.11 描绘了单位能量的态密度，在带边是 0，增长速率决定于电子和空穴的有效

质量. 在前面的表 2.1 中给出了 Si 和 GaAs 的 m_c和 m_v是适用于计算态密度的平均值.

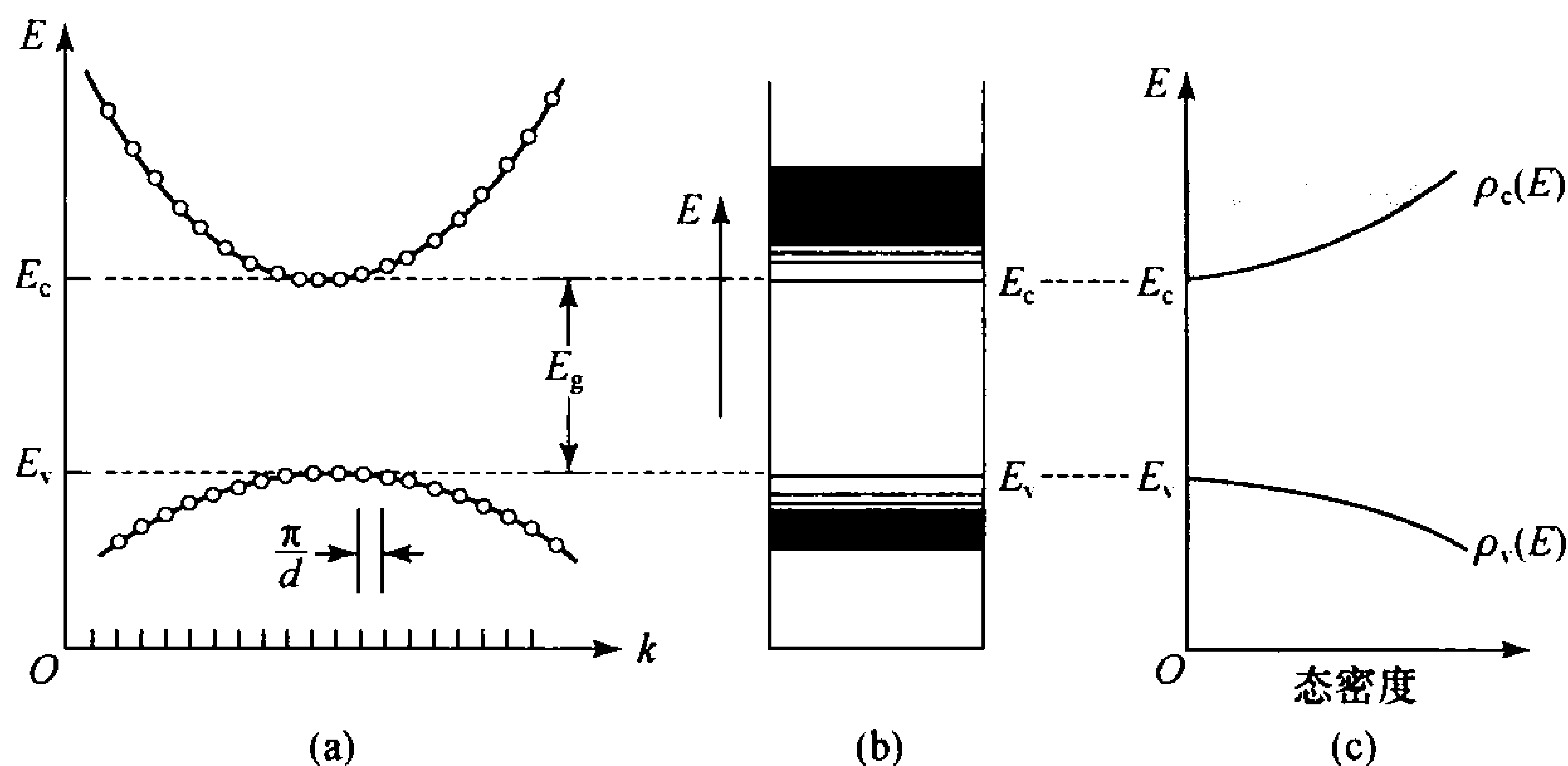


图 2.11　k_1 方向的 E-k 关系截面图(k_2,k_3 固定)(a)；允许的能级(对于所有的 k)(b)；导带和价带边的能量态密度函数(c)

2. 能级占有的概率

在没有热激发的情况下(T=0 K)，所有的电子遵循泡利不相容原理，占据最低的能级；价带被完全填满(不存在空穴)，导带完全空置(不存在电子). 当温度升高时，热激发使得有些电子离开价带到导带中去，在价带中留下空的量子态(空穴). 统计力学理论告诉我们，在热均衡温度 T 下，给定能量为 E 的能级被占优的概率满足费米函数

$$f(E)=\frac{1}{\exp[(E-E_f)/k_BT]+1},\qquad(2.2.6)$$

其中 k_B 为玻尔兹曼常数(T=300 K 时 k_BT=0.026 eV)；E_f 是一个常数，被称为费米能量或费米能级；这个方程又被称为费米-狄拉克分布. 能级 E 要么被占据[概率 $f(E)$]，要么空置[概率 $1-f(E)$]. 概率 $f(E)$和 $1-f(E)$根据式(2.2.6)由能量 E 的值决定. 概率 $f(E)$自己并不是一个概率密度分布，也不满足归一化；相反它仅仅指出一系列能级被占据的概率.

因为无论温度 T 是多少，$f(E_f)=1/2$，费米能级恒定表示被占据概率是 1/2 的能级(如果允许存在). 费米函数是一个 E 的单调递减函数(参见图 2.12). 在 T=0 K 时，如果 $E>E_f$，则 $f(E)=0$；否则，$f(E)=1$. 这表明导带中没有电子，价带中没有空穴. 费米能级 E_f 的显著性：温度为 0 K 时，它是能级被占据和未被占据的分界线. 因为 $f(E)$表示一个能级被电子占据的概率，$1-f(E)$则表示其未被占据的概率，也就是被空穴占据的概率(如果 E 是在价带中). 因此，对于能级 E：$f(E)$为能级被电子占据的概率，$1-f(E)$为能级被空穴占据的概率，这些函数关于费米能级对称.

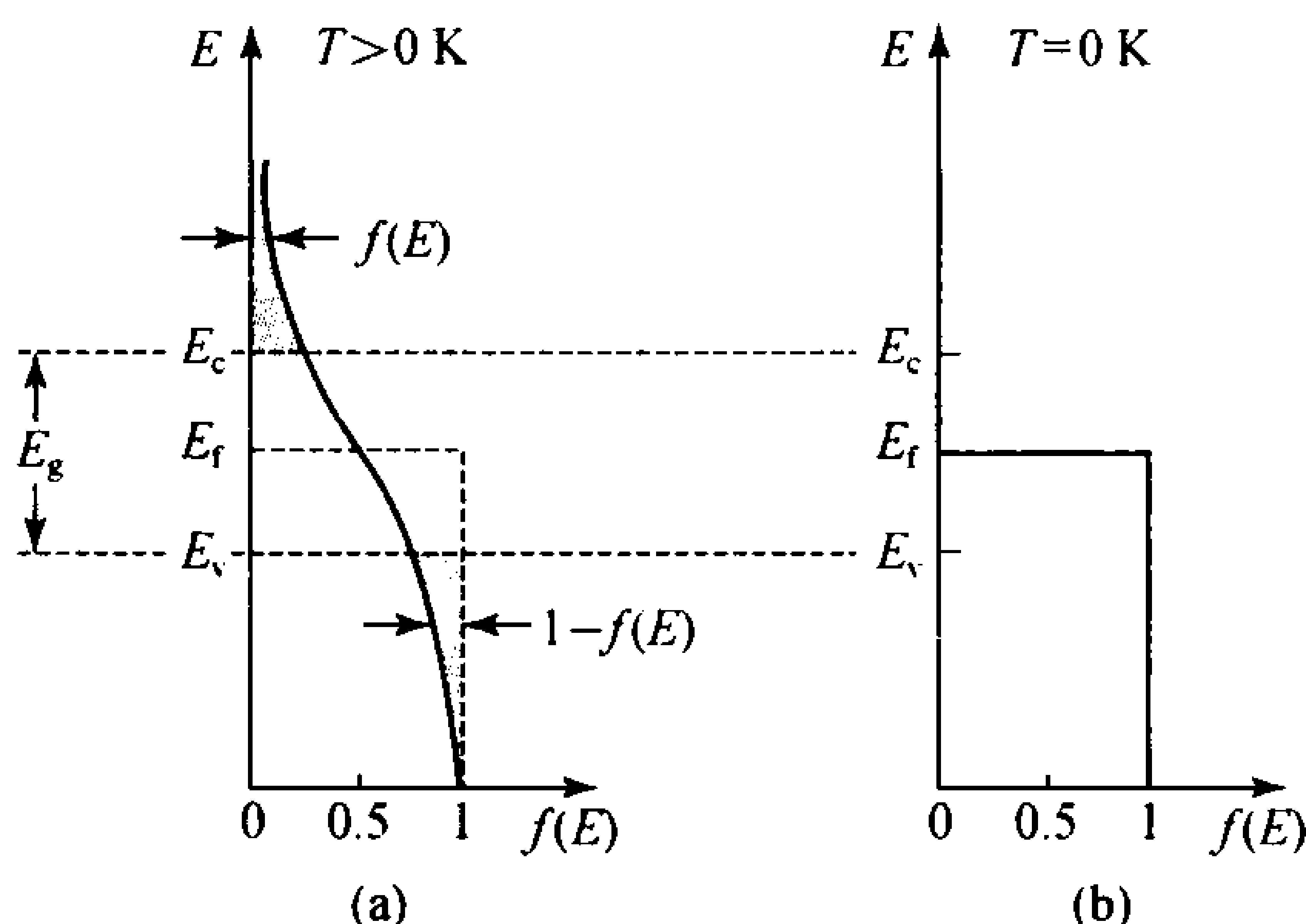


图 2.12　费米函数与能级占有概率的关系

当 $E-E_f \gg k_B T$，$f(E) \approx \exp[-(E-E_f)/k_B T]$时，费米函数在导带的高能量处随着能量上升以指数形式递减.此时费米函数正比于玻尔兹曼分布，它表示了部分原子被激发到给定能级后的指数依赖关系.换而言之，当 $E<E_f$ 且 $E_f-E \gg k_B T$，$1-f(E) \approx \exp[-(E_f-E)/k_B T]$时，也就是价带中能量远低于费米能级时，其被空穴占据的概率随着能量降低而指数性降低.

3. 热平衡条件下的载流子浓度

用 $n(E)\Delta E$ 和 $p(E)\Delta E$ 分别表示单位体积中能量从 E 到 $E+\Delta E$ 之间的电子和空穴数.$n(E)$和 $p(E)$的密度分布可以通过能量态密度和能级占有概率相乘而得到

$$n(E) = \rho_c(E)f(E), \quad p(E) = \rho_v(E)[1-f(E)]. \tag{2.2.7}$$

总的电子和空穴的浓度(单位体积的数量)n 和 p 由下述积分得到

$$n = \int_{E_c}^{\infty} n(E)\mathrm{d}E, \quad p = \int_{-\infty}^{E_v} p(E)\mathrm{d}E. \tag{2.2.8}$$

在本征(无掺杂)半导体中，任何温度下都有 $n=p$，因为热激发总是成对产生电子和空穴.费米能级必须是使得 $n=p$ 的能级.如果 $m_v=m_c$，那么 $n(E)$和 $p(E)$完全对称，因此 E_f 必须在带隙的正中间(参见图 2.13).在大多数本征半导体中费米能级的确是在带隙的正中间.

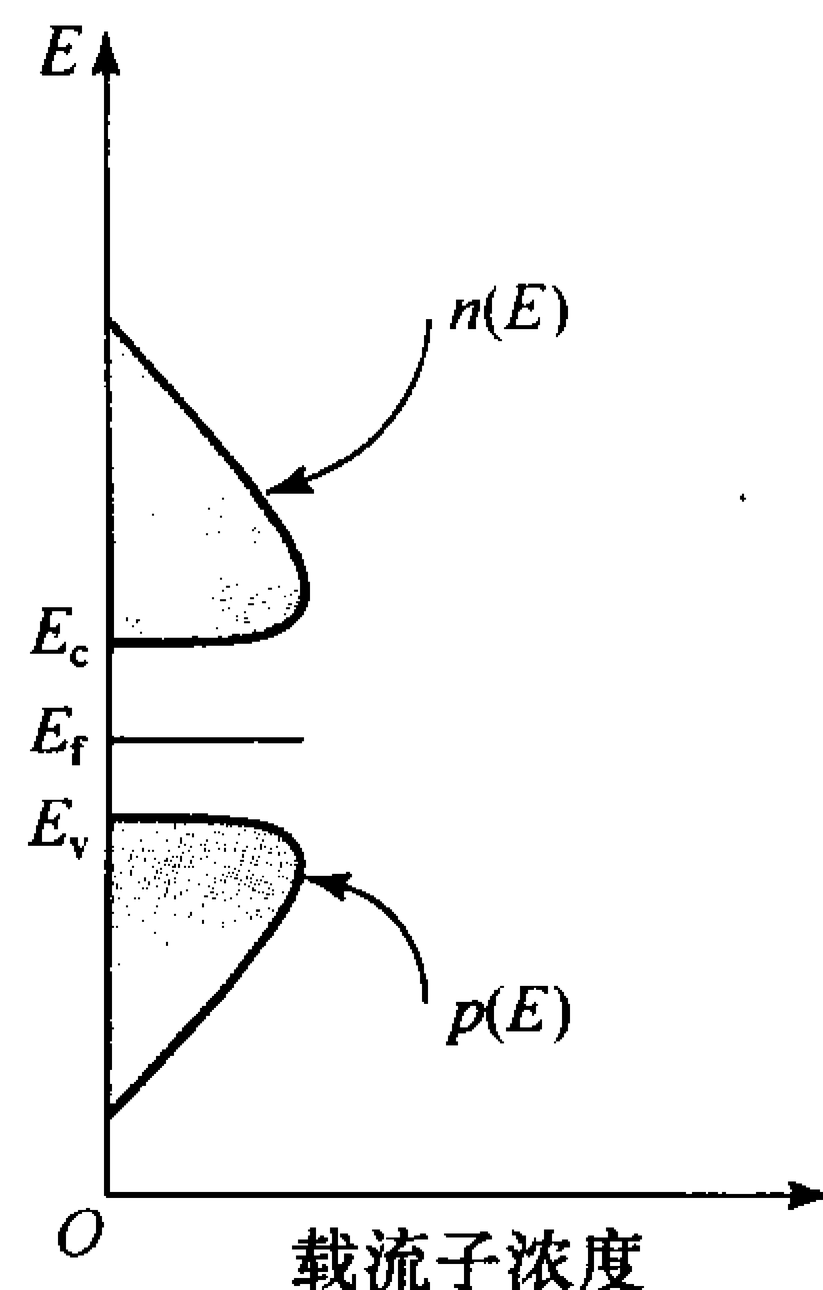


图 2.13　本征半导体中的电子、空穴浓度关于 E 的分布 $n(E)$和 $p(E)$

在图 2.14 和 2.15 中画出了热均衡时 n 型、p 型半导体的能带结构图、费米函数，以及电子、空穴浓度分布.施主电子占据着略低于导带底的能级 E_D，因此

容易被激发上导带. 如果 $E_D=0.01\ \text{eV}$,在室温下($k_BT=0.026\ \text{eV}$),大多数电子将被热激发至导带中. 因此,费米能级[$f(E_f)=1/2$] 将高于带隙的中间值. 在 p 型半导体中,受体占据着略高于价带顶的能级 E_A,因此费米能级低于带隙中间值. 我们的重点在于研究这些掺杂半导体材料中自由载流子的浓度. 这些材料都是电中性的,因此有关系 $n+N_A=p+N_D$,其中 N_A 和 N_D 分别是掺杂的施主和受主离子浓度.

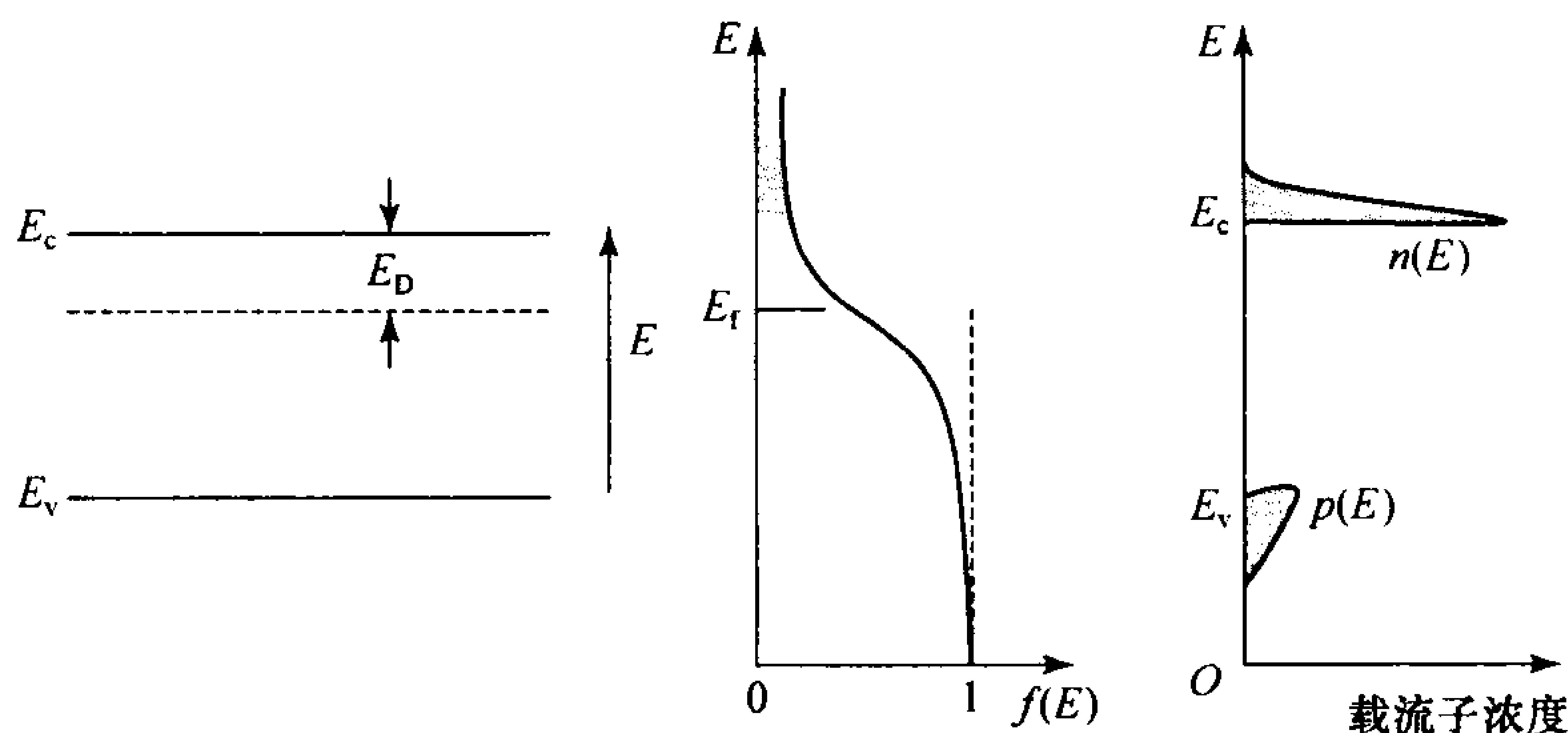


图 2.14 n 型半导体的能带结构图、费米函数和自由电子 $n(E)$、空穴 $p(E)$浓度分布

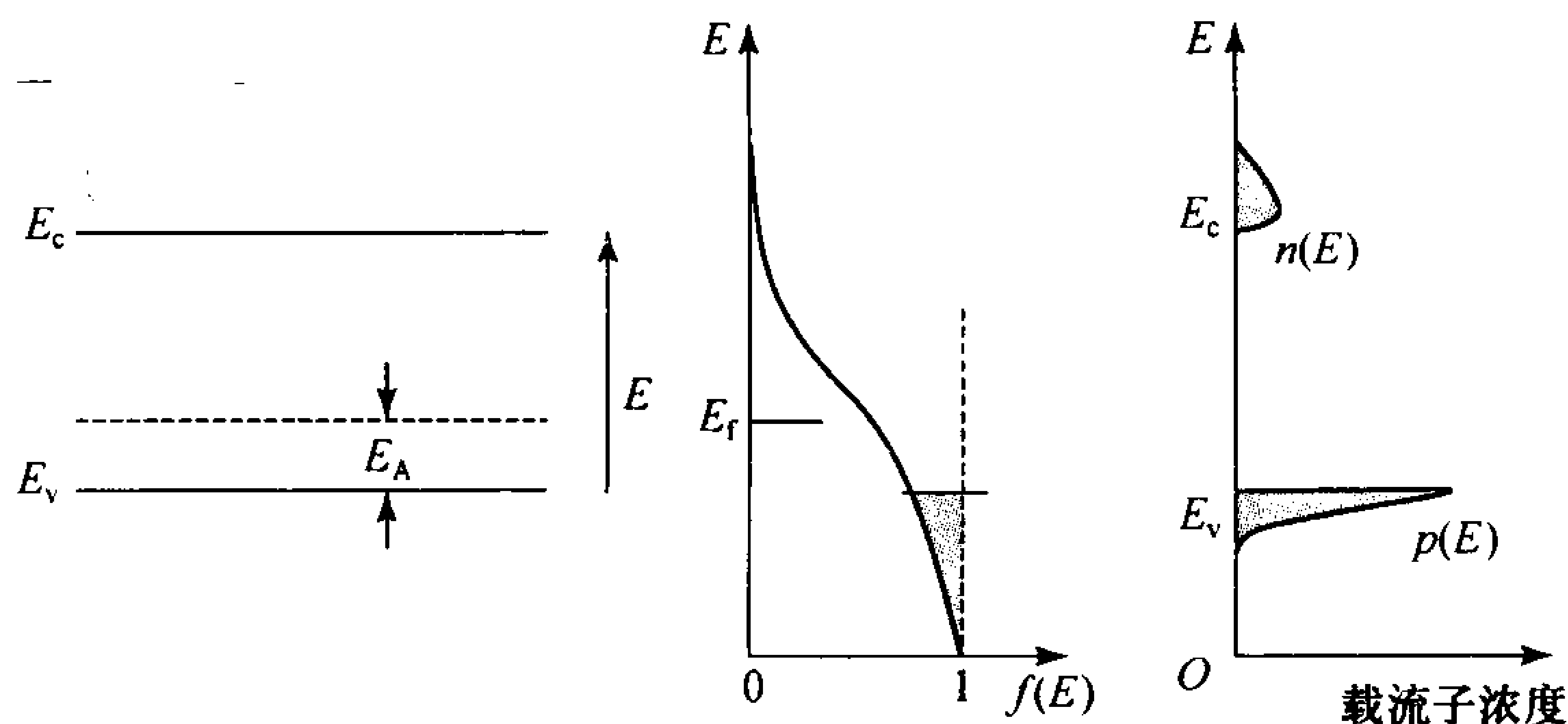


图 2.15 p 型半导体的能带结构图、费米函数和自由电子 $n(E)$、空穴 $p(E)$浓度分布

作为一个例子,我们来了解一下费米函数的指数近似. 当 $E-E_f\gg k_BT$ 时,费米函数 $f(E)$可以被近似为指数形式. 类似地,当 $E_f-E\gg k_BT$ 时,$1-f(E)$也可以近似为指数形式. 这些近似能够在费米能级存在于带隙中且和带边的距离是 k_BT 的几倍的情况下成立(室温下 $k_BT\approx0.026\ \text{eV}$,Si 和 GaAs 中的带隙 E_g 分别为 1.11 eV 和 1.42 eV). 将这些近似应用于本征和非本征半导体材料中,我们可以由式(2.2.8)获得如下关系式:

$$n=N_c\exp\left(-\frac{E_c-E_f}{k_BT}\right),\qquad(2.2.9\text{a})$$

$$p = N_v \exp\left(-\frac{E_f - E_v}{k_B T}\right), \tag{2.2.9b}$$

$$np = N_c N_v \exp\left(-\frac{E_g}{k_B T}\right), \tag{2.2.10a}$$

其中 $N_c = 2\left(\frac{2\pi m_c k_B T}{h^2}\right)^{3/2}$，$N_v = 2\left(\frac{2\pi m_v k_B T}{h^2}\right)^{3/2}$. 从上述关系式可以看出，当 $m_v = m_c$ 时，若 $n > p$ 则 E_f 更接近导带；反之若 $p > n$ 则 E_f 更接近价带.

在费米分布能够近似为指数形式的情况下，式(2.2.10a)表示下面的载流子浓度乘积：

$$np = 4\left(\frac{2\pi k_B T}{h^2}\right)^3 (m_c m_v)^{3/2} \exp\left(-\frac{E_g}{k_B T}\right), \tag{2.2.10b}$$

与费米能级 E_f 在带隙中的位置和掺杂浓度无关. 这种载流子浓度乘积的常数关系被称为质量作用定律.

对本征半导体而言，$n = p = n_i$. 与式(2.2.10a)联立得到本征半导体载流子浓度

$$n_i \approx (N_c N_v)^{1/2} \exp\left(-\frac{E_g}{2k_B T}\right), \tag{2.2.11}$$

此式表明本征半导体的电子和空穴浓度随着温度上升成指数级增加. 因此，质量作用定律也可以被写做如下形式：

$$np = n_i^2. \tag{2.2.12}$$

由于不同材料的带隙能量和有效质量不同，n_i 的取值也不相同. 室温下 Si 和 GaAs 的本征载流子浓度在表 2.2 中列举.

表 2.2　$T = 300$ K 时 Si 和 GaAs 的本征载流子浓度

	n_i/cm^{-3}
Si	1.5×10^{10}
GaAs	1.8×10^{6}

质量作用定律同样适用于决定载流子在掺杂半导体中的浓度. 例如，适度掺杂的 n 型半导体中自由电子的浓度 n 基本等于施主杂质的浓度 N_D. 利用质量作用定律，空穴浓度为 $p = n_i^2/N_D$. 知道了 n 和 p，我们就可以利用式(2.2.9)计算出费米能级. 只要费米能级在带隙中且距带边的能量是 $k_B T$ 的几倍，式(2.2.9)的近似就可以被直接运用.

如果费米能级在导带(或者价带)中，这样的材料被称为简并半导体. 这种情况下近似指数形式的费米分布函数不再成立，因此 $np \neq n_i^2$，我们只能通过数值算法的方法得到载流子浓度. 在重掺杂的情况下，施主(或受主)杂质的能带会和导带(或价带)合并形成所谓的带尾，这将使带隙显著变窄.

4. 准热平衡载流子浓度

上述关于能级占有概率和载流子浓度的讨论都是在热平衡条件下才成立. 如果热平衡的条件被破坏,上述理论并不成立. 当然可能存在这样的情况: 导带中的电子自身热平衡,价带中的空穴也自身热平衡,但是两者间并不是热平衡的. 例如,在外部电子流或光子流导致的带与带之间的高速跃迁,以至于很难维持导带和价带间的热平衡. 这样的情况被称为准热平衡,发生在带内跃迁的弛豫时间远远短于带间跃迁的弛豫时间. 一般地,能带内的弛豫时间小于 10^{-12} s,而电子空穴的辐射复合寿命约为 10^{-9} s.

在这样的情况下,应该给导带和价带引入各自的费米能级: E_{fc}和 E_{fv},被称为准费米能级(参见图 2.16). 当 E_{fc}和 E_{fv}分别处于导带和价带的深处时,电子和空穴的浓度可能会相当大.

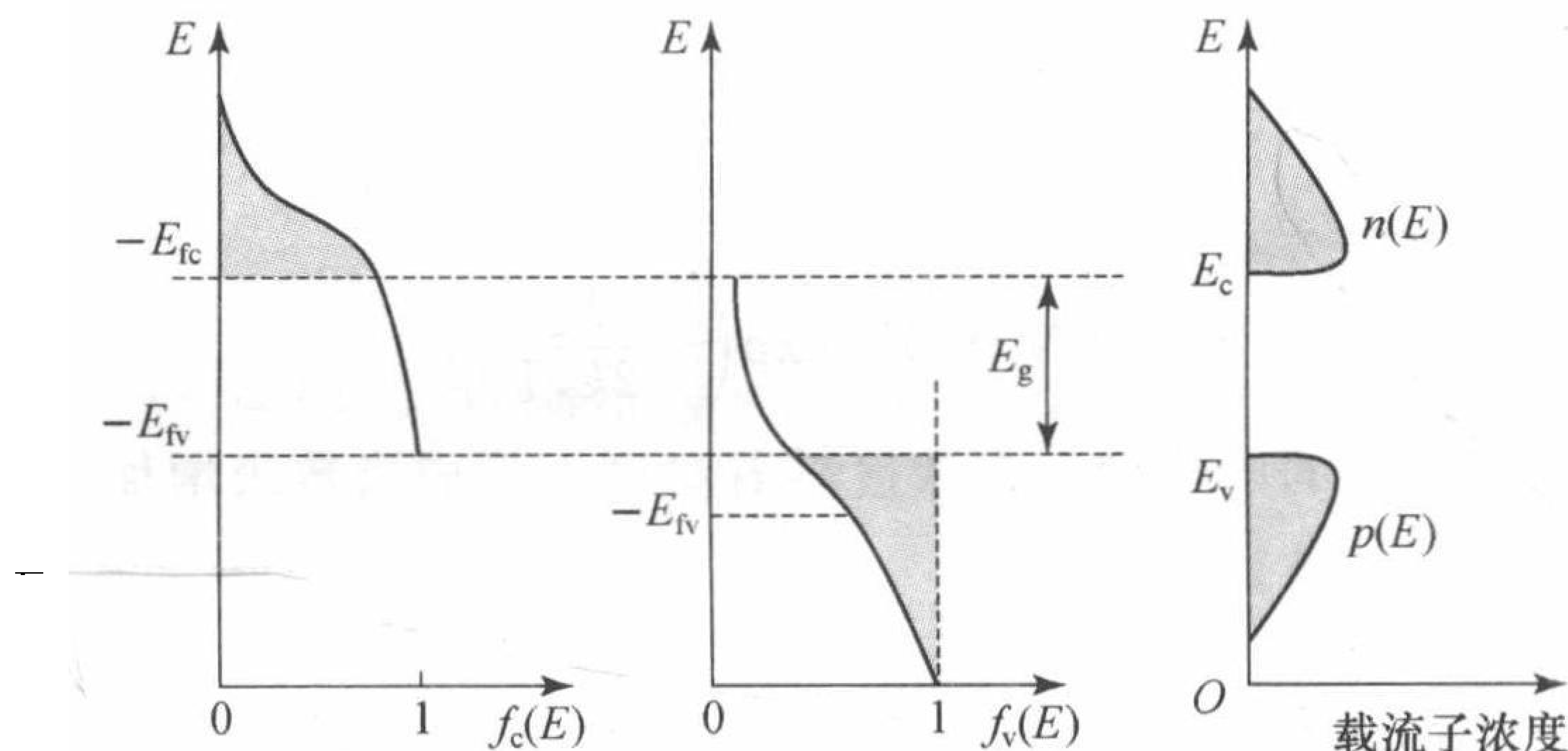


图 2.16 半导体的准费米能级

导带中能量为 E 的能级被占据的概率为 $f_c(E)$,它也是含有费米能级为 E_{fc}的费米分布函数. 价带中能量为 E 的能级被占据的概率为 $1-f_v(E)$,其中 $f_v(E)$ 是含有费米能级为 E_{fv}的费米分布函数. 电子和空穴的浓度为 $n(E)$和 $p(E)$,它们可能同时都比较大.

2.2.3 载流子的产生、复合和注入

载流子的产生、复合和注入是通过电子和空穴的同时参与而完成的. 因此,本小节描述的载流子行为规则既适用于电子,也可用来解释空穴.

1. 热平衡下载流子的产生与复合

价带电子被热激发到导带导致了电子空穴对的产生(参见图 2.17). 热平衡要求这一产生过程同时伴随着相反的退激发过程. 这一过程被称为电子空穴的复合,它的发生过程是一个自由电子从导带衰退下来填充价带空穴(参见图 2.17). 电子衰退的能量可以通过辐射光子的形式释放出来,这一过程称为辐射复合. 非辐

射复合可通过许多独立同时发生的过程而产生，这些过程包括了能量转变为晶格振荡(产生一个或多个声子)或者转变为另一个自由电子(俄歇过程).

复合过程也可能通过俘获或者缺陷中心间接产生(参见图2.18).它们是与杂质或者缺陷相关的能级，由晶粒边界、错位或者其他位于带隙中的晶格缺陷产生.一个杂质或者缺陷状态如果能够同时俘获电子与空穴，从而增加它们复合的可能性，那么该状态可以认为是一个复合中心.由杂质而增强的复合可以是辐射复合或者是非辐射复合.

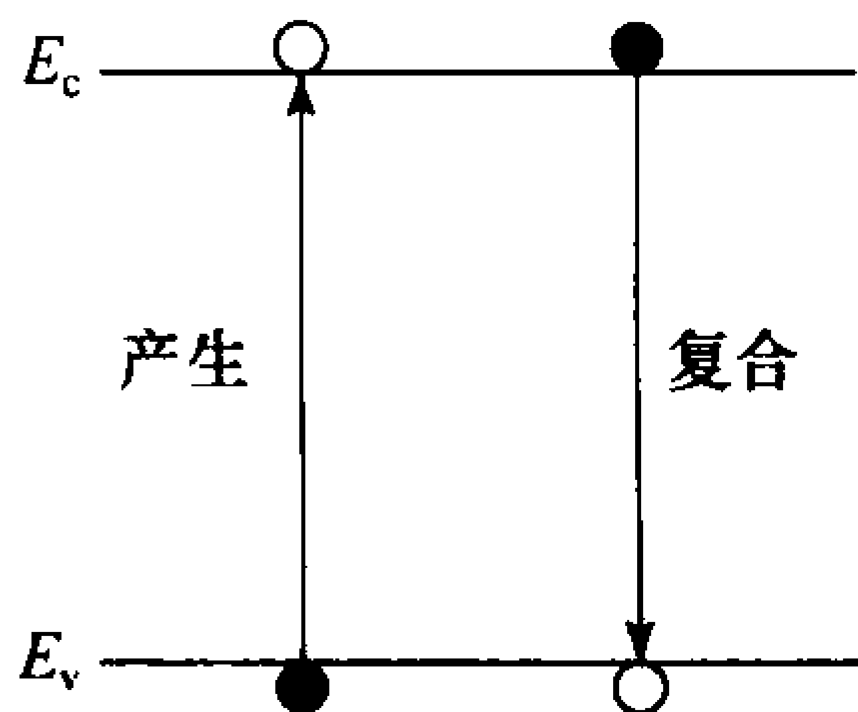


图2.17　电子-空穴对的产生与复合

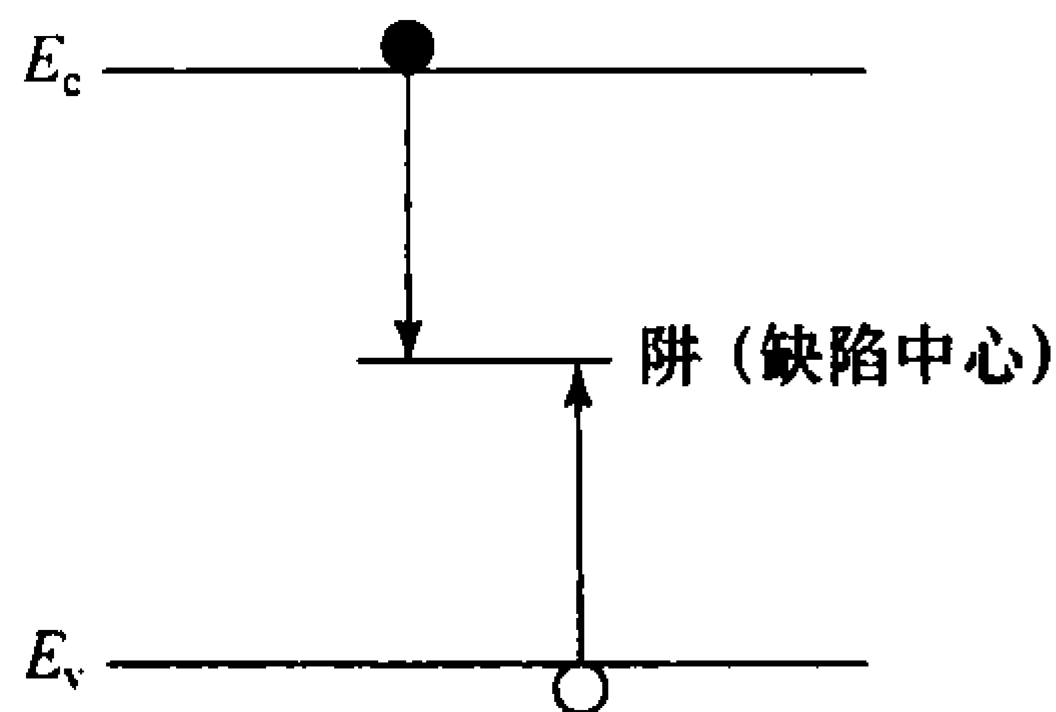


图2.18　由于阱(缺陷中心)而产生的电子-空穴对复合

因为复合的发生需要同时具备一个电子和一个空穴，所以复合的速率正比于电子和空穴的浓度的乘积，也就是说

$$\text{复合速率} = lnp, \tag{2.2.13}$$

上式中 l(单位为 cm^3/s)是一个变量，该变量依赖于包括构成成分、缺陷和温度在内的材料特征，也弱弱地依赖于材料的掺杂.

当载流子的产生和复合保持平衡时，电子和空穴的浓度分别为 n_0 和 p_0.在稳定的状态下，复合速率必须等于产生速率.如果 G_0 是在一定温度下热载流子(电子空穴对)的产生速率，那么在热平衡条件下，

$$G_0 = ln_0 p_0.$$

电子、空穴浓度的乘积 $n_0 p_0 = G_0/l$，不管在n型、p型，还是本征材料中都是近似相等的.因此 $n_i^2 = G_0/l$，直接导出了质量作用定律 $n_0 p_0 = n_i^2$.可见，质量作用定律是在热平衡条件下保证载流子产生和复合平衡的必然结果.

2. 热平衡下载流子的注入

在热平衡情况下，一种载流子浓度分别为 n_0 和 p_0 的半导体，其产生速率等于复合速率，即 $G_0 = ln_0 p_0$.现在利用外部注入机制(非热力学)使半导体产生额外的电子空穴对，其速率为稳定的 R(单位体积电子空穴对数量).当达到一个新的稳定状态时，载流子浓度分别为 $n = n_0 + \Delta n$ 和 $p = p_0 + \Delta p$.很明显，由于电子和空穴是成对出现，所以无论如何都满足 $\Delta n = \Delta p$.由于新的产生率等于复合率，我们得到

$$G_0 + R = lnp. \tag{2.2.14}$$

将 $G_0=ln_0p_0$ 代入式(2.2.14),得到

$$R=l(np-n_0p_0)=l(n_0\Delta n+p_0\Delta n+\Delta n^2)=l\Delta n(n_0+p_0+\Delta n).$$

上式可以写成如下形式:

$$R=\Delta n/\tau, \tag{2.2.15}$$

其中

$$\tau=\frac{1}{l(n_0+p_0+\Delta n)}. \tag{2.2.16}$$

由于注入速率满足 $\Delta n\ll n_0+p_0$,所以

$$\tau=\frac{1}{l(n_0+p_0)}, \tag{2.2.17}$$

此时的 τ 被称为过剩载流子复合寿命.

在 n 型材料中,$n_0\gg p_0$,所以复合寿命 $\tau\approx 1/(ln_0)$ 与电子浓度成反比.类似地,对于 p 型半导体来说,$p_0\gg n_0$,我们可以得到 $\tau=1/(lp_0)$.上述简单方程在俘获起着重要作用的过程中不适用.

参数 τ 可以看做是注入的过剩电子-空穴对的复合寿命.如果注意到注入载流子浓度满足下面的速率方程,那么上述对于 τ 的解释就很容易理解了:

$$\frac{\mathrm{d}(\Delta n)}{\mathrm{d}t}=R-\frac{\Delta n}{\tau}.$$

在稳定状态下 $\mathrm{d}(\Delta n)/\mathrm{d}t=0$,于是就得到式(2.2.15).如果在 t_0 时刻瞬间将注入源移去(R 变为 0),那么 Δn 将随着时间常数 τ 以指数形式衰减,即 $\Delta n(t)=\Delta n(t_n)\exp[-(t-t_n)/\tau]$.在另一方面,如果是强注入,由式(2.2.16)可看出,τ 本身就是 Δn 的函数,因此速率方程不再是线性的,衰减时间也不再是指数的.

如果注入速率 R 已知,则稳定状态的注入浓度可以由下式决定:

$$\Delta n=R\tau. \tag{2.2.18}$$

总的浓度则可以由 $n=n_0+\Delta n$ 和 $p=p_0+\Delta n$ 决定.而且如果准平衡假设成立,那么式(2.2.8)可以用来计算准费米能级.准平衡状态并没有与上述分析的产生率与复合率之间的平衡所矛盾,它仅仅要求带内平衡时间相比复合时间 τ 来说很短即可.在研究半导体发光二极管和半导体激光器过程中,上述分析将会被证明是十分有用的.这些器件是通过载流子注入的方式来增强光发射的.

3. 内量子效率

半导体材料的内量子效率 η_i 被定义为产生辐射的电子空穴复合速率与总的(辐射的与非辐射的之和)复合速率之比.这一参数非常重要因为它决定了半导体材料中光产生的效率.总的复合速率由式(2.2.13)决定.如果参数 l 被分为辐射部分和非辐射部分,即 $l=l_r+l_{nr}$,那么内量子效率为

$$\eta_i = \frac{l_r}{l} = \frac{l_r}{l_r + l_{nr}}. \quad (2.2.19)$$

由于 τ 反比于 l[见式(2.2.17)]，所以内量子效率也可以写成复合寿命的形式. 定义辐射和非辐射寿命分别为 τ_r 和 τ_{nr}，故

$$\frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_{nr}}. \quad (2.2.20)$$

由于$\frac{l_r}{l} = \frac{1}{\tau_r} \Big/ \frac{1}{\tau}$，内部量子效率可写成

$$\eta_i = \frac{\tau}{\tau_r} = \frac{\tau_{nr}}{\tau_r + \tau_{nr}}. \quad (2.2.21)$$

辐射复合寿命 τ_r 决定了光子吸收和发射的速率，它的值由载流子浓度和材料参数 l_r 决定. 对于小的到中等的注入速率有

$$\tau_r \approx \frac{1}{l_r(n_0 + p_0)}, \quad (2.2.22)$$

与式(2.2.17)相符. 非辐射复合寿命 τ_{nr} 被一个相似的方程所决定. 然而，如果非辐射复合是由带隙中的缺陷中心所导致的，那么 τ_{nr} 对于这些缺陷中心浓度的敏感程度要高于对电子空穴浓度的敏感程度，即缺陷中心浓度对于 τ_{nr} 影响更大.

表 2.3 提供了 Si 和 GaAs 的复合速率及寿命的近似值. l_r，τ_r(假设温度为 300 K，载流子浓度为 $n_0 = 10^{17}\ cm^{-3}$ 的 n 型材料)、τ_{nr}(假设缺陷中心浓度为 $10^{15}\ cm^{-3}$)、τ 以及内量子效率 η_i 的数量级也在表中给出.

表 2.3　Si 和 GaAs 的辐射复合速率 τ_r、复合寿命、内量子效率 η_i 的近似值

	$l_r/(cm^3 \cdot s^{-1})$	τ_r	τ_{nr}/ns	τ/ns	η_i
Si	10^{-15}	10 ms	100	100	10^{-5}
GaAs	10^{-10}	100 ns	100	50	0.5

Si 的辐射寿命在数量级上大于它的总寿命，其原因在原理上可以解释为 Si 有间接带隙. 这导致了较小的内部量子效率. 另一方面，对于 GaAs 来说，衰减很大程度上依靠辐射跃迁(具有直接带隙)，所以其内部量子效率较大. 因此 GaAs 和其他直接禁带材料对于制造光发射结构是较为适用的，Si 和其他间接禁带材料就不是非常适用了.

2.2.4　p-n 结

1. 同质结

同一半导体材料的不同掺杂区域之间的结合称为同质结. 不同半导体材料之间的结合称为异质结. 同质结中的一个重要例子是 p-n 结. p-n 结是 p 型和 n

型材料所形成的同质结. 在电子学中 p-n 结可以被制作成整流器、逻辑门、稳压器(稳压二极管)和调谐器(变容二极管);在光电子学中它可以用做发光二极管(LED)、激光二极管、光电探测器和太阳能电池.

在掺杂半导体材料中, p 型材料有许多空穴(多数载流子)和很少的自由电子(少数载流子);n 型材料有许多自由电子和少量空穴. 两种载流子都朝各个方向进行连续随机热运动. 其载流子浓度及能级结构如图 2.19 所示.

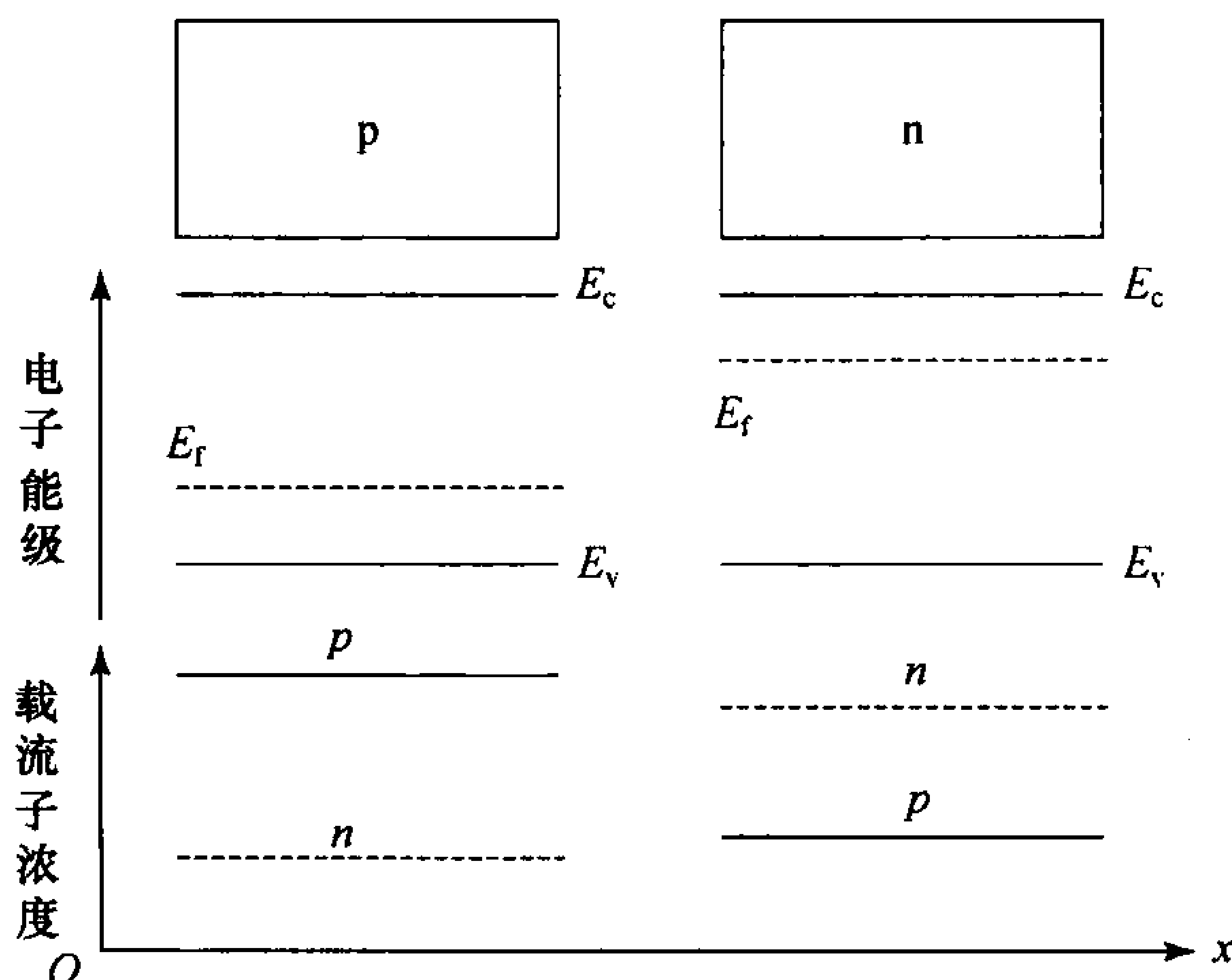


图 2.19 接触前一个 p 型半导体和一个 n 型半导体的能级以及载流子浓度

当两种材料金相接触时,形成 p-n 结(图 2.20),下列事件就发生了. 在图 2.20 中耗尽层、能带图、自由电子浓度 $n(x)$ 以及空穴的浓度 $p(x)$(指数形式)都以位置 x 的函数形式被表示出来. 内建势相当于能量 eV_0,其中 e 是电子电量.

① 电子和空穴从高浓度区域向低浓度区域扩散. 这一扩散过程不能无限地持续下去,然而,它却破坏了两区域间的电荷平衡.

② 自由载流子在结两边的一个狭窄区域内基本上消失. 这一区域被称为耗尽层. 它仅含有固定电荷(n 型区域的正离子和 p 型区域的负离子). 两区域耗尽层的厚度反比于该区域掺杂物的浓度.

③ 固定电荷在耗尽层中产生了由 n 型区域指向 p 型区域的电场. 这一内建场阻碍了自由载流子在结区域的进一步扩散.

④ 平衡状态的建立导致在耗尽层两端产生内建势, n 型区域的电势高于 p 型区域.

⑤ 内建势使得 n 型区域的电子相比 p 型区域有较低的势能. 结果,能带就如图 2.20 那样弯曲了. 在热平衡状态下整个结构只有单一的费米函数,所以 n 型和 p 型区域的费米能级必须一致.

⑥ 没有穿过结的净电流. 电子和空穴的扩散和漂移电流各自抵消.

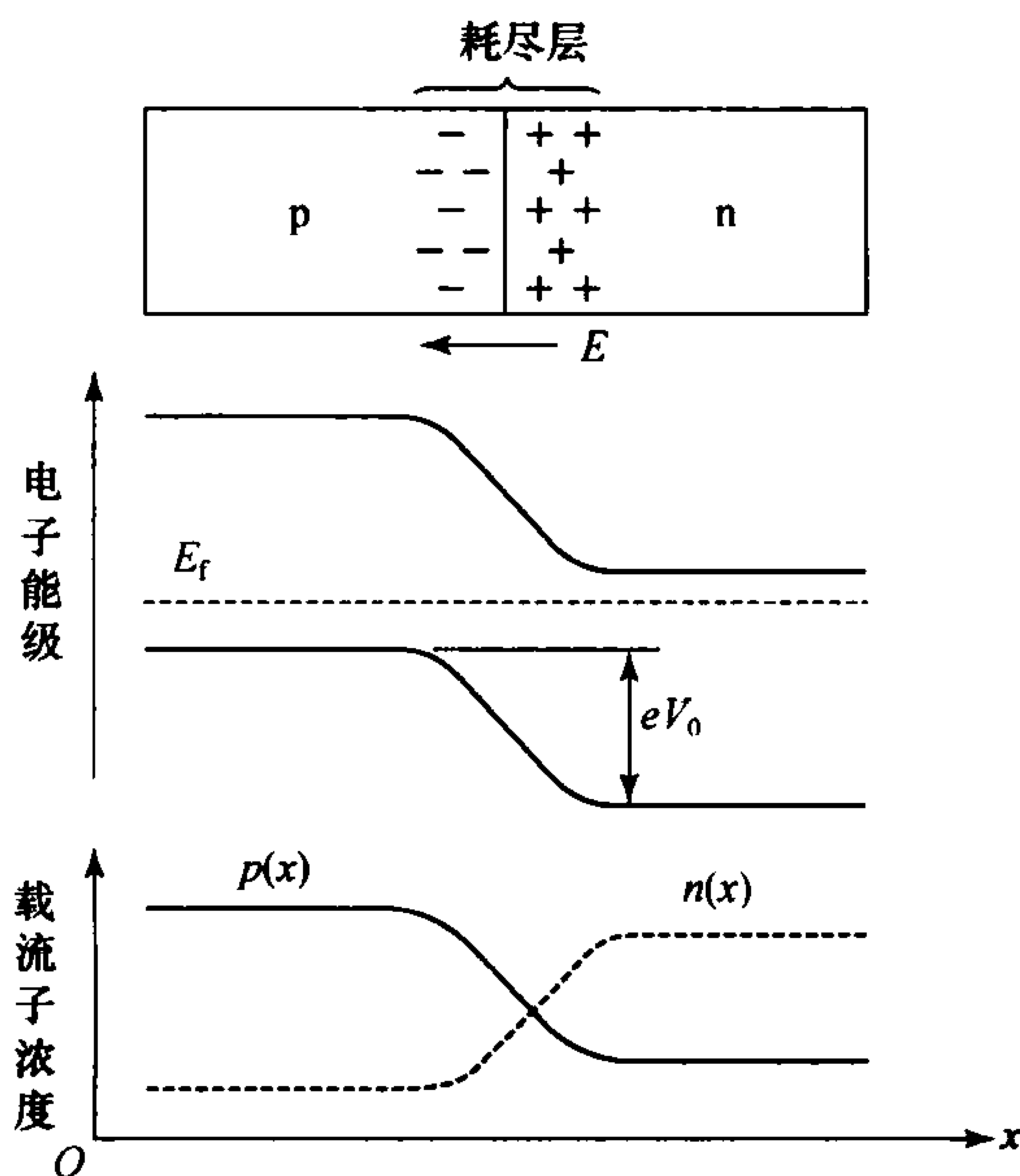


图 2.20　$T>0$ K 时热平衡状态下的一个 p-n 结

(1) 偏置结.

一个外部电势的作用将改变 p 型和 n 型区域间的电势差. 这将转而改变多数载流子的流量,因此这种结可以被用做"门". 如果通过对于 p 型区域施加正向电压 V 使得结获得正向偏置(图 2.21),那么它的电势相对于 n 型区域来

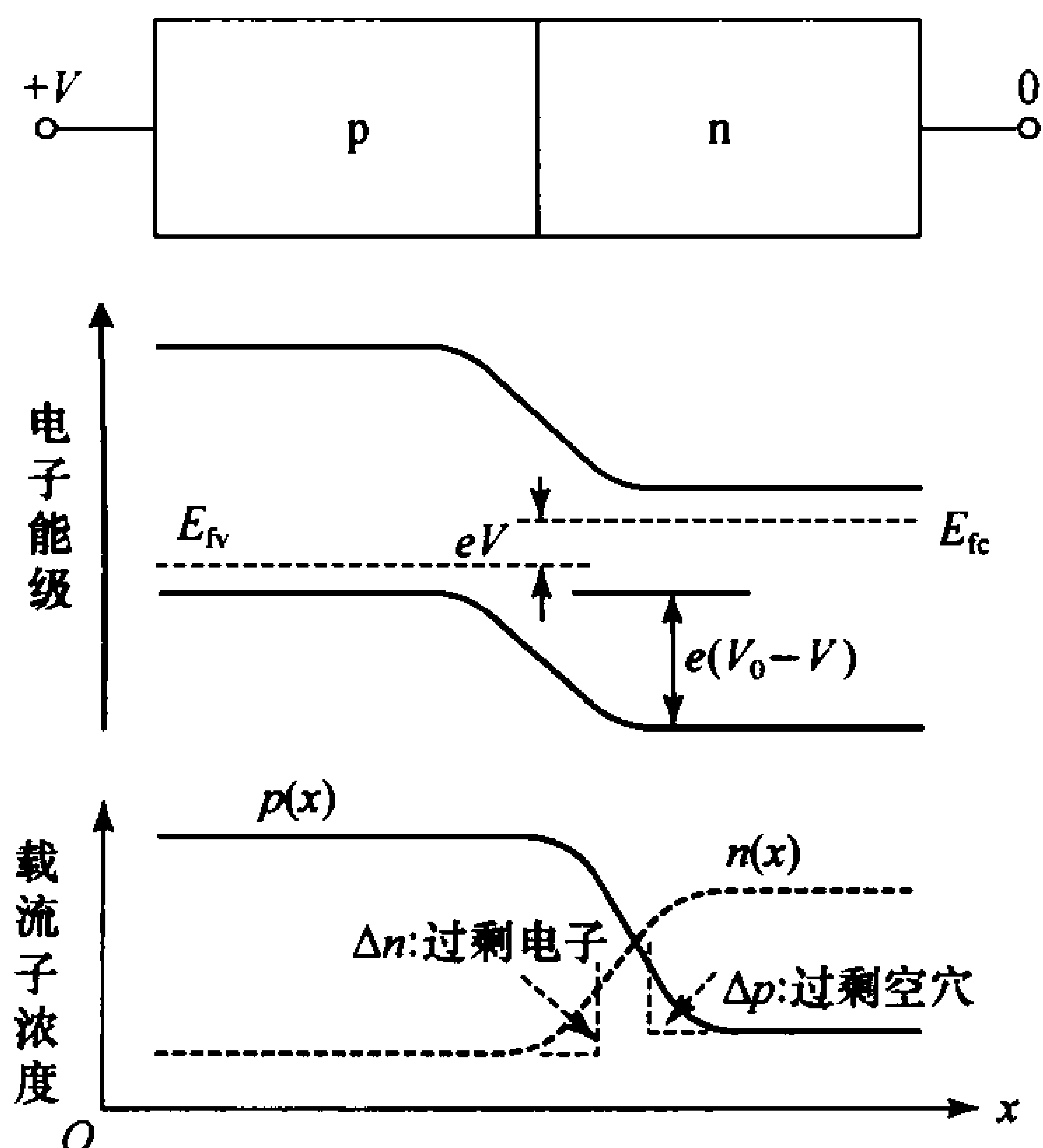


图 2.21　正向偏置的 p-n 结能带图和载流子浓度

说变大了，因此其电场方向与内建势的电场方向相反．外部偏置电压的存在导致了偏离平衡以及 n 型区域与 p 型区域的费米能级的不一致，同时也导致了耗尽层费米能级的不一致．耗尽层中两个费米能级的存在，E_{fc}和 E_{fv}，代表了准平衡状态．

正向偏置的净效应为势垒高度的降低，降低数量为 eV．多数载流子电流最终按指数规律 $\exp(eV/k_BT)$增长，因此净电流变为 $i=i_s\exp(eV/k_BT)-i_s$，此处 i_s 为一常数．

过剩的多数空穴和电子分别进入 n 型和 p 型区域，它们将变为少数载流子并和该区域的多数载流子复合．因此它们的浓度将随着离结的距离增大而减小，如图 2.21 所示．这一过程称为少数载流子注入．

如果在 p 型区域施加负电压 V 使得结反向偏置，那么势垒高度将增加 eV，这将阻碍多数载流子的流动．相应的电流被乘以了指数因子 $\exp(eV/k_BT)$，此处 V 是负数，也就是说电流减少了．净电流是 $i=i_s\exp(eV/k_BT)-i_s$，因此当 $|V|\gg k_BT/e$，大小约为 i_s 的小电流将流向相反方向．因此 p-n 结作为一个理想二极管拥有 i-V 特性：$i=i_s[\exp(eV/k_BT)-1]$，如图 2.22(c)所描述．

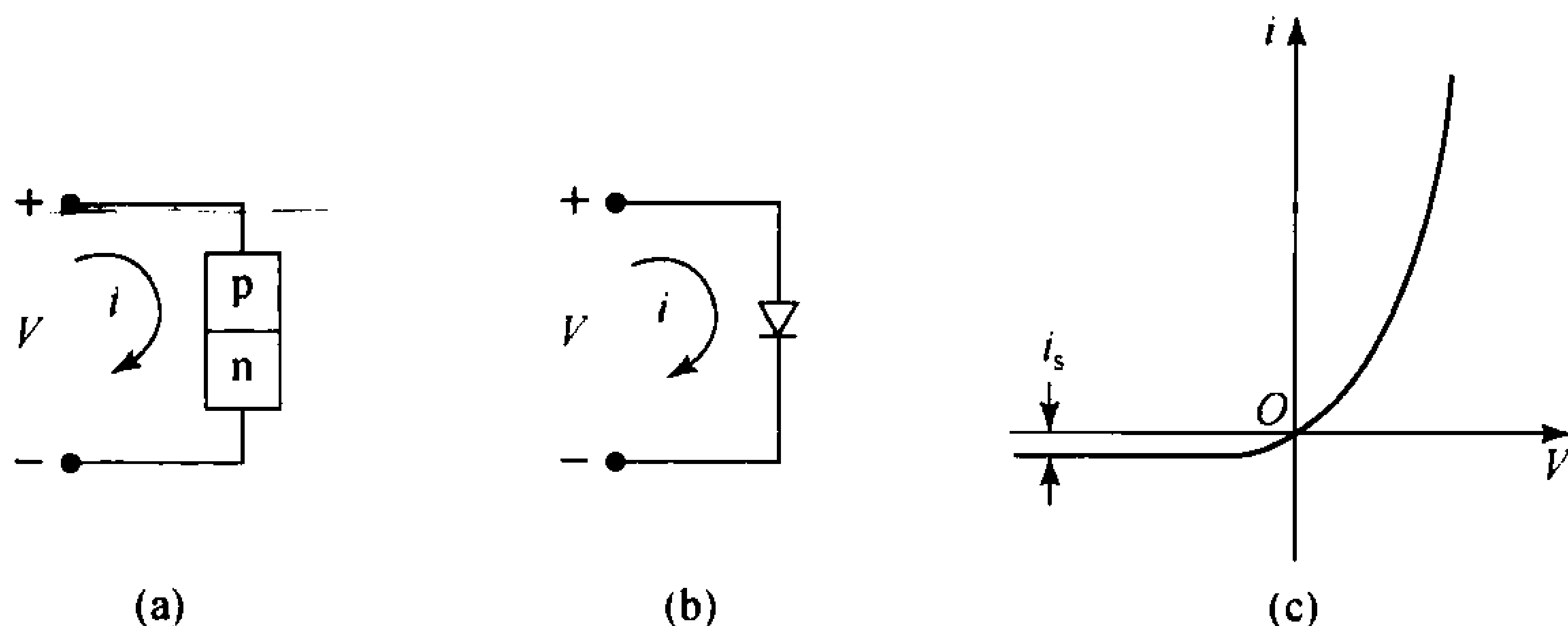


图 2.22 一个 p-n 结的电压和电流(a)，p-n 结二极管的电路表示(b)，理想 p-n 结二极管的电流-电压特征(c)

p-n 结的动态特性可以通过一组决定电子空穴扩散、(在内建势和外部电场的影响下的)漂移和复合过程的方程来得到．这些对于了解二极管的工作频率是十分重要的．其理论模型可以由两个电容(一个结电容和一个扩散电容)并联一个理想二极管来获得．结电容代表了当外加电压改变时存储在耗尽层内的固定正负电荷改变所需的时间．耗尽层厚度 l 正比于$(V_0-V)^{1/2}$，因此它在反向偏置情况下(V 为负数)增大，在正向偏置情况下(V 为正数)减小．因此结电容 $C=\epsilon A/l$(此处 A 是结面积)反比于$(V_n-V)^{1/2}$．反向偏置二极管的结电容小于正向偏置二极管的结电容，因此 RC 反应时间较短．C 依赖于 V 的性质可用于制作可变电压电容器(变容二极管)．

正向偏置二极管中少数载流子的注入可以用扩散电容来描述，该电容依赖于少数载流子寿命和工作电流.

(2) p-i-n 结.

p-i-n 结是在 p 型区域和 n 型区域之间插入一层本征(或轻掺杂)半导体材料(参见图 2.23). 因为耗尽层延伸到结两边的距离反比于掺杂浓度，所以 p-i 结的耗尽层渗透入 i 区域比较深. 类似地，i-n 结的耗尽层较深地延伸进 i 区域. 因此，p-i-n 就像一个耗尽层扩展到整个本征区域的 p-n 结. 热平衡时 p-i-n 结的电子能量，固定电荷密度以及电场如图 2.23 所示. 应用一个拥有较宽耗尽层的结的好处之一是它有较小的结电容以及较快的反应速度. 因此，相比 p-n 结，p-i-n 结更具有潜力被用作半导体光电二极管. 较宽的耗尽层同时也使得更多的入射光可以被捕获，因此增加了光电测控的效率.

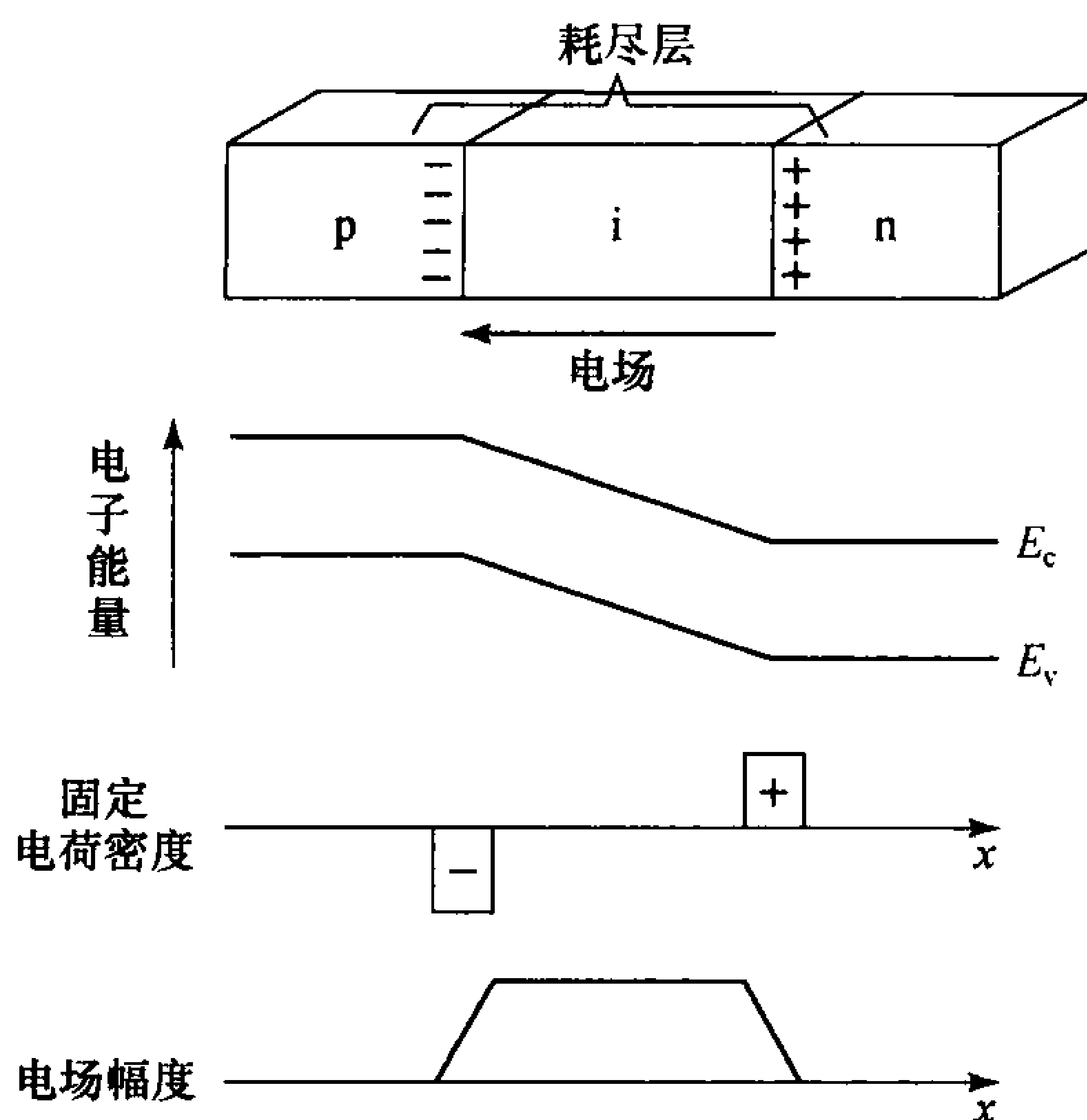


图 2.23　热平衡状态下 p-i-n 二极管的电子能量，固定电荷密度和电场大小

2. 异质结

不同半导体材料之间的结被称为异质结. 异质结可以被用来制作新型双极型和场效应晶体管，也可以在光源及光探测器中得到应用. 它们使得电子和光电子设备性能有了长足的进步. 特别是在光子学中，将不同的半导体材料并列排在一起可以在如下几方面产生优势：

① 不同禁带宽度的材料之间的结可以在能带图上产生局部的跃变. 势能的不连续可以产生一个阻碍，使得某些载流子无法进入不希望它们进入的区域. 这一性质可以在 p-n 结中得以应用，比如用来减少少数载流子的电流比例，从而增加注入效率(参见图 2.24).

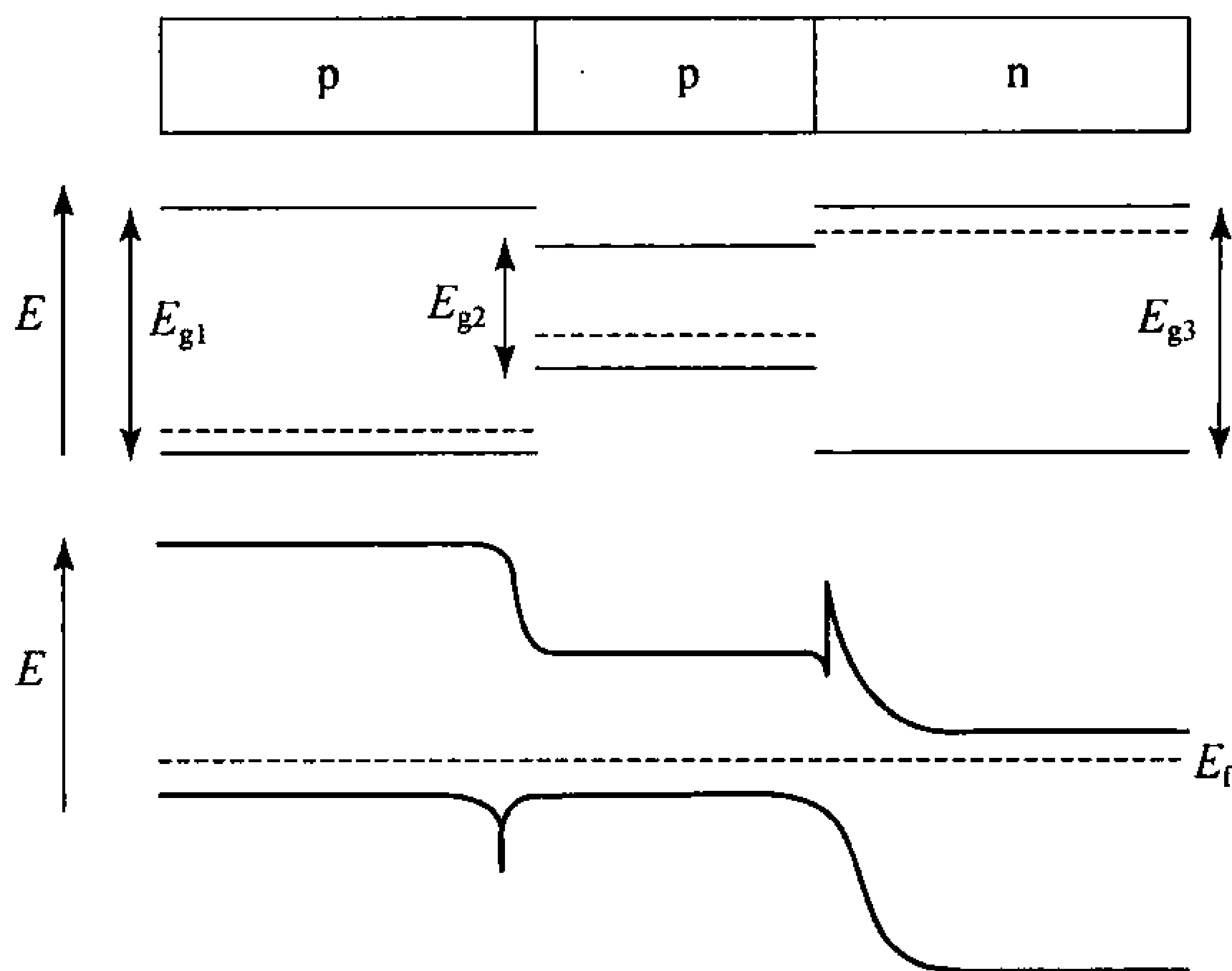


图 2.24 p-p-n 双异质结结构

如图 2.24 所示，中间层的禁带宽度比两边要窄. 平衡时，由于费米能级要保持一直线，所以 p-p 结的导带边缘迅速下降，p-n 结的价带边缘也迅速下降. 导带能量差与价带能量差之比称为能带失调值. 当器件处于正向偏置，这些跃变将成为势垒从而限制注入的少子. 比如，从 n 区域注入的电子将会被阻挡使其无法扩散并越过 p-p 结的势垒. 类似地，从 p 区域注入的空穴也会被阻挡使其无法扩散并跃过 p-n 结的势垒. 因此这种双异质结迫使电子和空穴占据很窄的共有区域. 这对于注入式激光二极管有效地工作是非常必要的.

② 两个异质结所导致的能带图的不连续，对于将载流子限制在希望的空间区域中是很有用的. 比如，一层窄禁带宽度的材料可以被两层宽禁带材料夹在中间，如图 2.24 所示的 p-p-n 结构(它包含了一个 p-p 异质结和一个 p-n 异质结). 这种双异质结被广泛应用于制作激光二极管.

③ 异质结对于产生可在特定区域加速载流子的能带不连续性是十分有效的. 在一个多层雪崩光电二极管中，这些突然增加的动能可以有选择性地增加载流子碰撞电离的概率.

④ 不同带隙类型(直接与间接)的半导体可以在同一器件中被用于选择光发射的区域，因为只有直接带隙半导体才能有效地发光.

⑤ 不同带隙的半导体可以在同一器件中被用于选择光的吸收区域. 能带宽度大于入射光子能量的半导体材料是透明的，其作用与“窗户层”相同.

⑥ 不同折射系数材料组成的异质结可被用于制造光波导，它可以限制并引导光子的传播.

3. 量子阱和超晶格

半导体薄层材料的异质结构可以通过外延生长的方式来获得，也就是说，通过一些技术使得一种半导体材料通过晶格匹配的方式覆盖于另一种之上. 这些技

术包括分子束外延法(molecular beam epitaxy, MBE),液相外延法(liquid phase epitaxy, LPE)以及气相外延法(vapour phase epitaxy, VPE).气相外延法的一种特殊形式是金属有机化学气相沉积(metal organic chemical vapour deposition, MOCVD). MBE是利用所需元素的分子束撞击一块在高真空环境下精心准备好的基片;LPE是利用含有所需元素的冷却饱和溶液与基片接触;MOCVD则是将气体注入一个反应装置中.不同薄层的组分和掺杂可以通过控制分子的到达速率和基片表面的温度来确定.这些薄层可以薄得与分子层一样厚.

当薄层厚度与德布罗意波长可比甚至更小时(比如在GaAs中德布罗意波长约为50 nm),适用于体型半导体材料的能量-动量关系将不再准确.在微纳光子学中,通常涉及三种重要的结构:量子阱、量子线和量子点.适用于这些结构的能量-动量关系将在下面进行讨论.

(1) 量子阱.

量子阱是一种双异质结结构,包含一层超薄的材料(厚度不超过50 nm),其带宽小于周围其他材料[参见图2.25(a)]. Si材料的薄层被SiO_2材料包围,GaAs材料的薄层被AlGaAs材料包围就是例子.这种"三明治"的结构形成了导带和价带的矩形势阱,电子和空穴则被限制在该势阱中:电子在导带势阱中,空穴在价带势阱中.一个足够深的势阱可以近似被认为是无限深势阱.

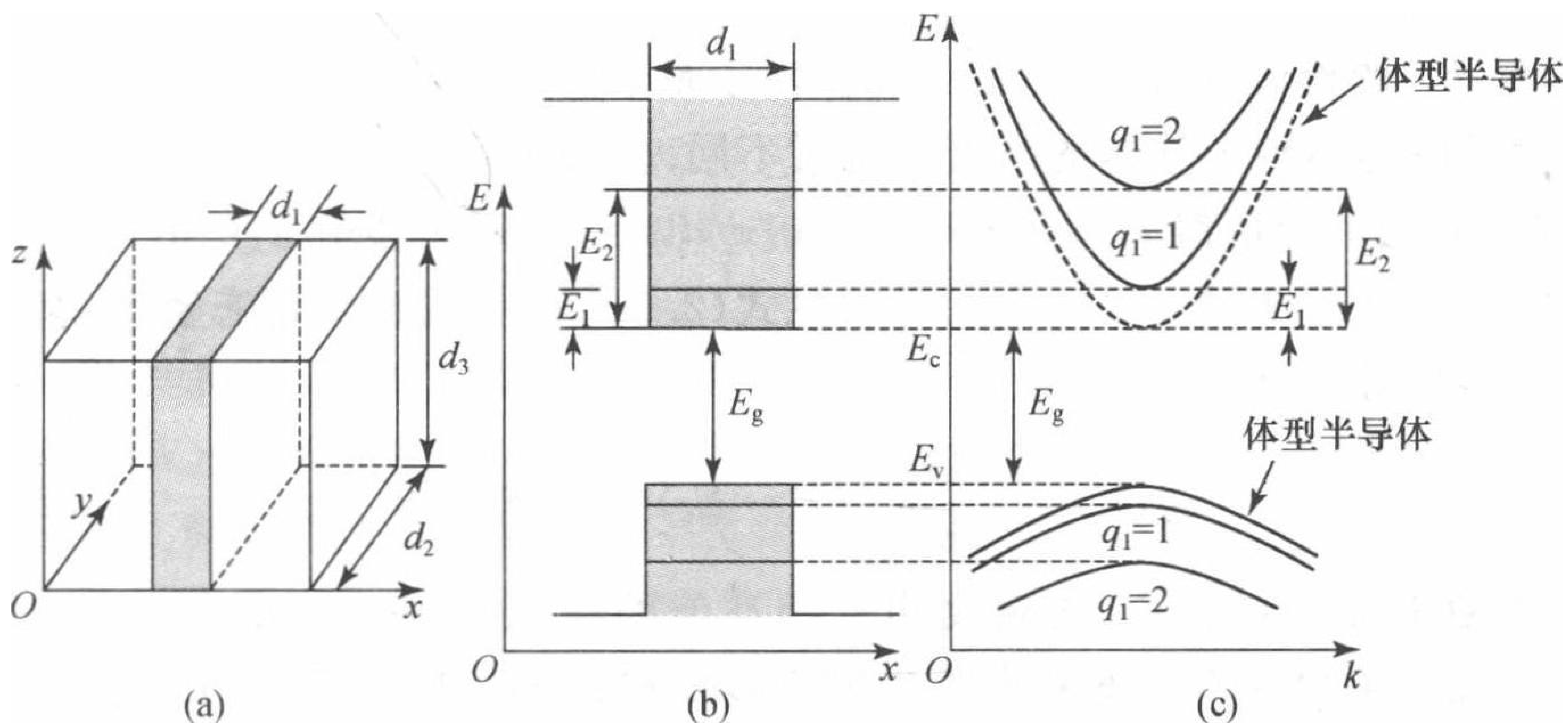


图2.25　量子阱的几何结构(a),量子阱中电子空穴的能带图(b),k_2或k_3方向上E-k关系的横截面(c);次能带由它们的量子数$q_1=1,2,\cdots$所标出.体型半导体的E-k关系由虚线显示

一个质量为m(电子有效质量m_c,空穴有效质量m_v)的颗粒,如果被限制在一个一维无限矩形势阱之中,假定该势阱全宽为d,那么它的能级为E_q可以通过解定态薛定谔方程得到

$$E_q=\frac{\hbar^2(q\pi/d)^2}{2m},\quad q=1,2,3,\cdots. \tag{2.2.23}$$

作为例子，我们知道在一个无限深的宽度为 $d=10$ nm 的 GaAs 势阱中($m_c=0.07m_0$，m_0 为电子静止质量)，因而，允许的电子能级为 $E_q=54$ meV，216 meV，486 meV，…($T=300$ K 时，$k_BT=26$ meV). 势阱宽度越小，相邻能级间的分隔就越大.

在图 2.25 所示的量子阱结构中，电子(和空穴)在 x 方向上被限制在 d_1(势阱的厚度)的一段距离中. 然而，它们在限制层平面的另两个方向上却延伸了很多(d_2，$d_3\gg d_1$). 因此在 Oyz 平面上，它们的行为就和在体型半导体中一样. 与该结构相对应的能量-动量关系可以被表述为

$$E=E_c+\frac{\hbar^2k_1^2}{2m_c}+\frac{\hbar^2k_2^2}{2m_c}+\frac{\hbar^2k_3^2}{2m_c},$$

式中 $k_1=q_1\pi/d_1$，$k_2=q_2\pi/d_2$，$k_3=q_3\pi/d_3$，且 $q_1,q_2,q_3=1,2,3,\cdots$. 由于 $d_1\ll d_2$，d_3，所以 k_1 的值非常离散，然而 k_2 和 k_3 却是只有微小空间离散性的值，它们可以被近似看做连续体. 因此量子阱导带中电子动量-能量关系为

$$E=E_c+E_{q_1}+\frac{\hbar^2k^2}{2m_c},\quad q_1=1,2,3,\cdots,\qquad(2.2.24)$$

其中 k 是 Oyz 平面内二维矢量 $\boldsymbol{k}=(k_2,k_3)$的数值. 每一个量子数 q_1 对应于一个最低能量为 $E_c+E_{q_1}$ 的次能带. 价带之中也有类似的关系.

体型半导体的能量-动量关系由式(2.2.1)给出，其中 k 是三维波矢 $\boldsymbol{k}=(k_1,k_2,k_3)$的模值 $|\boldsymbol{k}|$. 唯一的区别是，对于量子阱来说，k_1 取非常分离的值. 因此，量子阱与体型半导体的态密度有所不同. 体型半导体态密度决定于它的三维波矢的三个分量值 $k_1=q_1\pi/d$，$k_2=q_2\pi/d$，$k_3=q_3\pi/d$ 其中 $d_1=d_2=d_3=d$. 结果是每单位体积有 $\rho(k)=k^2/\pi^2$[参见式(2.2.3)]，这导致导带态密度[参见式(2.2.4)和图 2.26]为

$$\rho_c(E)=\frac{\sqrt{2}m_c^{3/2}}{\pi^2\hbar^3}(E-E_c)^{1/2},\quad E>0.\qquad(2.2.25)$$

而在量子阱结构中，态密度由二维波矢的模值(k_2，k_3)而得到. 对于每个量子数 q_1，在 Oyz 平面态密度为每单位面积有 $\rho(k)=k/\pi$，因此为每单位体积 $k/(\pi d_1)$. 密度 $\rho_c(E)$和 $\rho(k)$的关系为 $\rho_c(E)\mathrm{d}E=\rho(k)\mathrm{d}k=[k/(\pi d_1)]\mathrm{d}k$. 最终，利用 E-k 关系(2.2.24)我们得到 $\mathrm{d}E/\mathrm{d}k=\hbar^2k/m_c$，从而

$$\rho_c(E)=\begin{cases}\dfrac{m_c}{\pi\hbar^2d_1}, & E>E_c+E_{q_1},\\ 0, & E<E_c+E_{q_1},\end{cases}\quad q_1=1,2,3,\cdots.\qquad(2.2.26)$$

因此对于每个量子数 q_1，当 $E>E_c+E_{q_1}$ 时单位体积态密度为定值. 总的态密度是所有 q_1 值对应态密度之和，因此它的分布如图 2.26 所示那样是阶梯形的. 阶梯的每个台阶等同于一个不同量子数 q_1 并且可以被视做导带中的一个次能带

(图 2.26). 这些次能带底随着量子数的增长而迅速增长. 将 $E=E_c+E_{q_1}$ 代入式(2.2.25)中并利用式(2.2.23),我们发现在 $E=E_c+E_{q_1}$ 时,量子阱的态密度与体型半导体的态密度一样. 价带中的态密度也有类似的阶梯形分布.

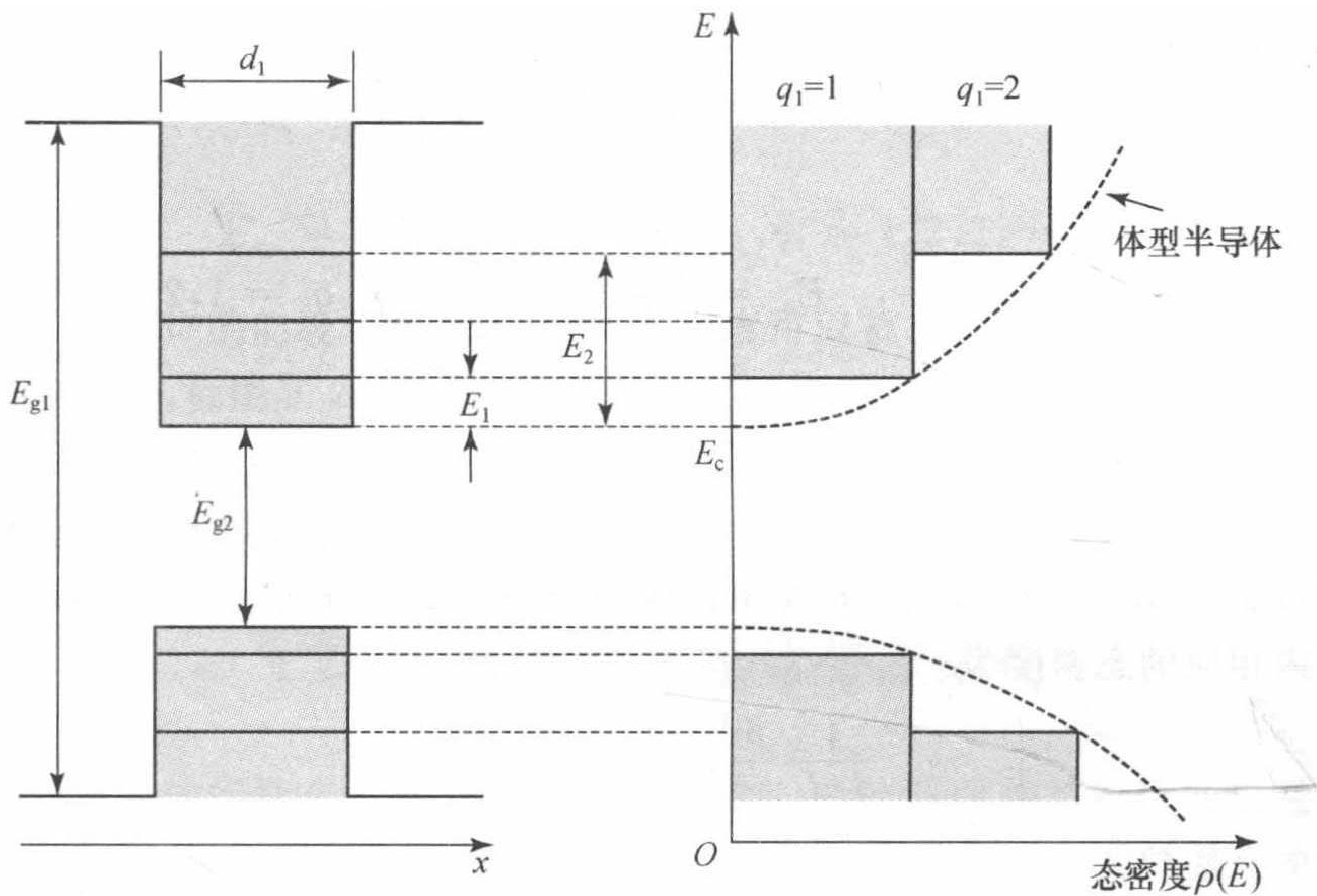


图 2.26　量子阱结构的态密度(实线)和体型半导体的态密度(虚线)

相比于体型半导体,量子阱结构在它的最低允许导带能级和最高允许价带能级体现出大量的态密度. 这一性质对于材料的光学性质具有很大的影响.

(2) 多量子阱和超晶格.

不同半导体材料互相交替组成的多层结构叫做多量子阱(multi quantum well, MQW)结构(参见图 2.27). 现在有许多不同的方法可以用来制造出能级带隙随着位置不同而不同的多量子阱结构. 如果相邻势阱间的能量势垒足够薄,使得电子能够轻易隧穿势垒,那么离散的能级可以通过扩展而进入微小的能带. 在这种情况下,多量子阱结构也叫做超晶格结构. 多量子阱结构被用于激光和光学探测器,以及非线性光学元件中. 一种典型的 MQW 结构包含厚度约为 10 nm 的材料 100 层,每一层拥有 40 个原子面,因此结构的总厚度约为 1 μm. 这样的结构在 MBE 机器中生长需要约 1 h 的时间.

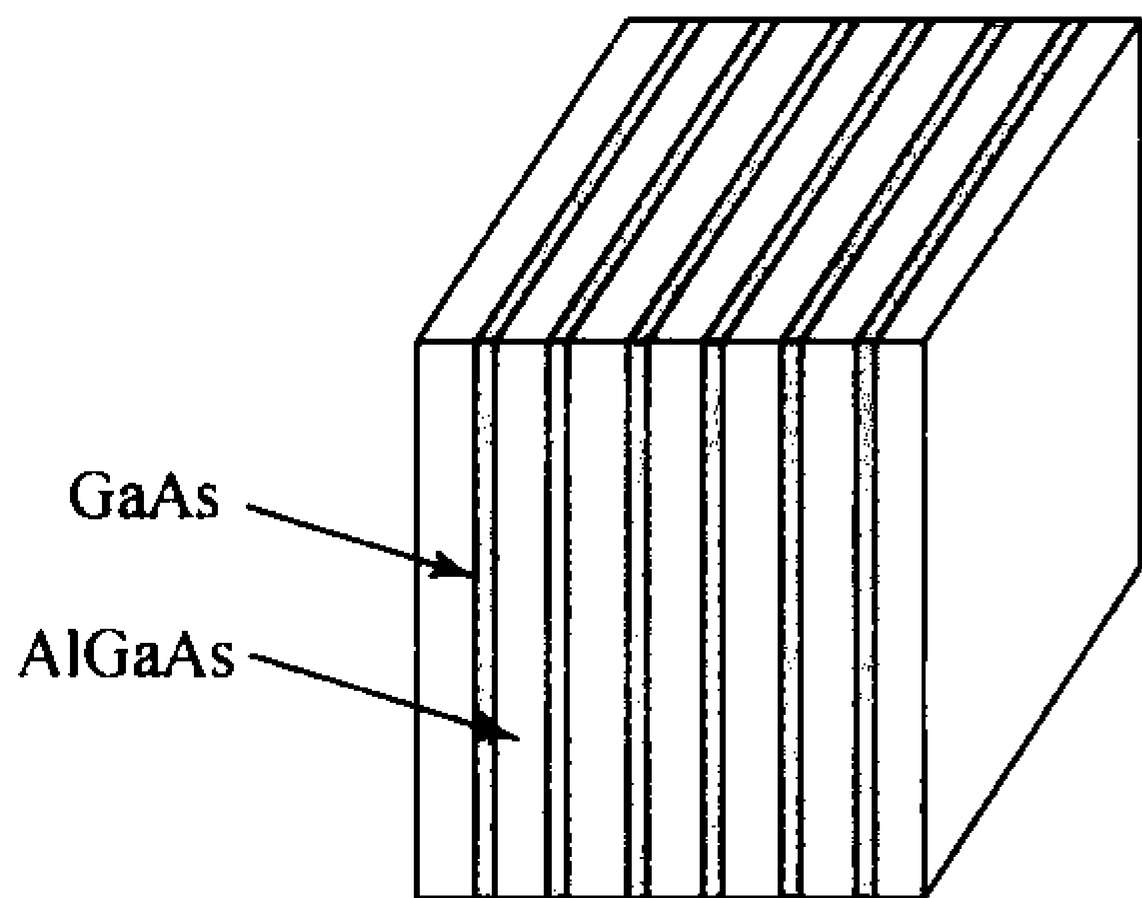


图 2.27　通过交替的 AlGaAs 层和 GaAs 层制造一个多层量子阱结构

(3) 量子线和量子点.

量子线结构(参见图 2.28)为拥有矩形横截面的细线形式的半导体材料,且周围被较宽带隙的材料包围.量子线有势阱的作用,可以牢牢地将电子(空穴)束缚在两个方向(x,y).假设横截面面积为 d_1d_2,则导带中的能量-动量关系为

$$E=E_c+E_{q_1}+E_{q_2}+\frac{\hbar^2k^2}{2m_c},\tag{2.2.27}$$

式中

$$E_{q_1}=\frac{\hbar^2(q_1\pi/d_1)^2}{2m_c},\quad E_{q_2}=\frac{\hbar^2(q_2\pi/d_2)^2}{2m_c},\quad q_1,q_2=1,2,3,\cdots,\tag{2.2.28}$$

且 k 是 z 方向的波矢分量(沿着量子线的轴).

每一对量子数(q_1,q_2)都与一个次能带相关联,该次能带的态密度为每单位长度量子线 $\rho(k)=1/\pi$,也就是每单位体积 $1/(\pi d_1d_2)$.作为能量的函数,其单位体积内相应的态密度为

$$\rho_c(E)=\begin{cases}\dfrac{\dfrac{1}{d_1d_2}\dfrac{m_c^{1/2}}{\sqrt{2}\pi\hbar}}{(E-E_c-E_{q_1}-E_{q_2})^{1/2}}, & E>E_c+E_{q_1}+E_{q_2},\\ 0, & E\leqslant E_c+E_{q_1}+E_{q_2},\\ q_1,q_2=1,2,3,\cdots. & \end{cases}\tag{2.2.29}$$

它们是能量的递减函数,如图 2.28(c)所示.这些量子线中的次能带比量子阱中的更窄.

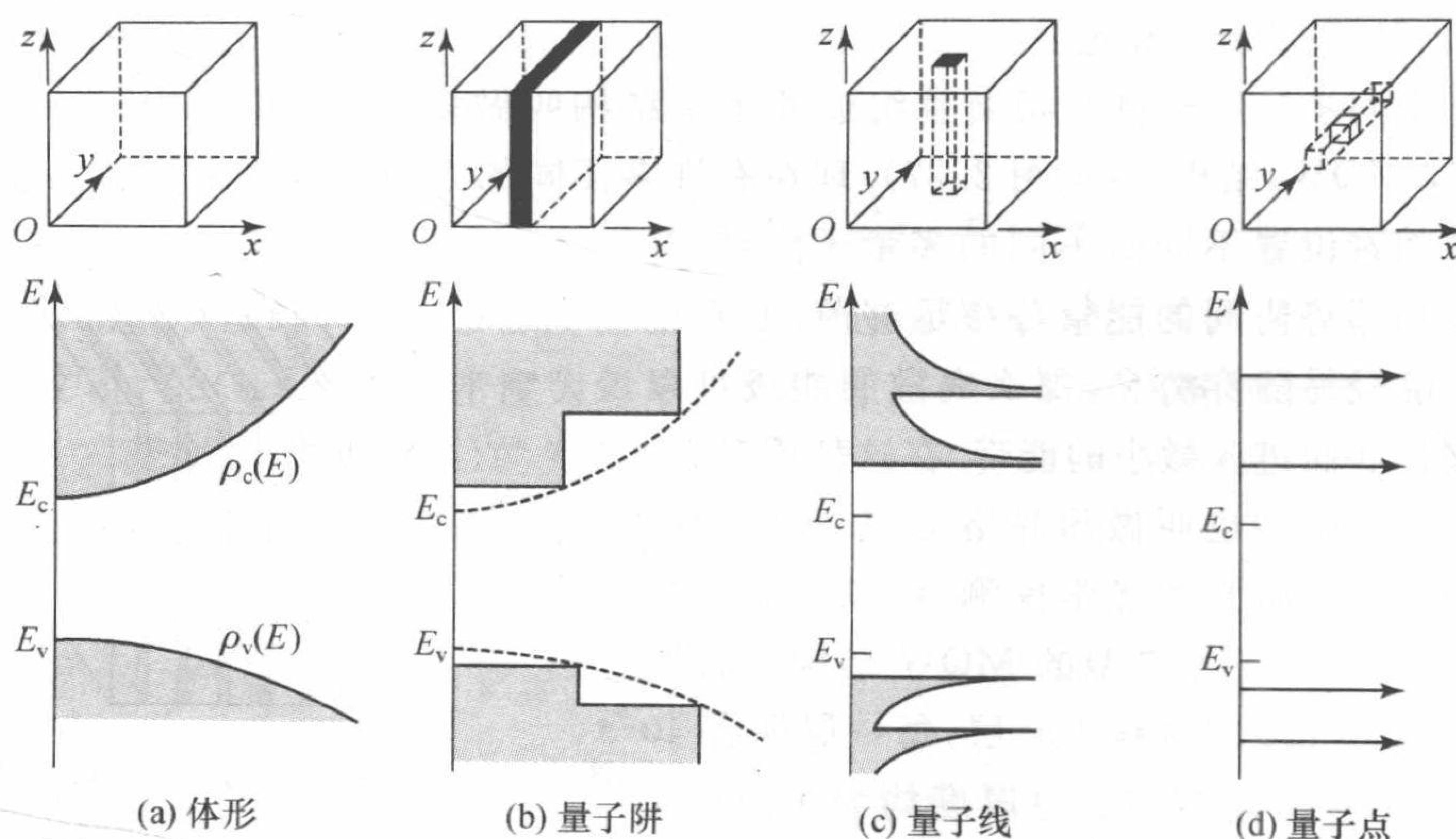


图 2.28　不同限制形状的态密度.导带和价带分裂为重叠的次能带,由于电子运动在更多维度上被限制,所以次能带会依次变窄

在量子点结构中,电子在全部三维方向中都被牢牢限制在了一个体积为 $d_1d_2d_3$ 的盒子中,因此能量被量子化为

$$E = E_c + E_{q_1} + E_{q_2} + E_{q_3},$$

式中

$$E_{q_1} = \frac{\hbar^2(q_1\pi/d_1)^2}{2m_c}, \quad E_{q_2} = \frac{\hbar^2(q_2\pi/d_2)^2}{2m_c},$$

$$E_{q_3} = \frac{\hbar^2(q_3\pi/d_3)^2}{2m_c}, \quad q_1, q_2, q_3 = 1, 2, 3, \cdots. \tag{2.2.30}$$

由于被允许的能级非常离散,因此态密度可以被一系列在允许能量处的脉冲函数所表示,如图 2.28(d)所示.量子点常被叫做人工原子.虽然量子点内可能拥有数万个强烈相互作用的自然原子,然而,原则上我们依旧可以通过适当设计任意选择量子点的离散能级.

§2.3 硅基光子晶体带隙结构

同固体晶体所具有的周期微结构特征相似,光子晶体的结构特征也是周期性的;与固体晶体由原子和电子组成周期的电子势垒不同,光子晶体是介电材料在波长尺度上其折射率按周期排列的人工电磁晶体.1987 年,美国 Bell 实验室的 Yablonovitch[2] 和美国普林斯顿大学的 John[3] 分别在研究如何抑制自发辐射和无序电介质材料中的光子局域时,各自独立地提出了“光子带隙”(photonic bandgap, PBG)和“光子局域”(localization of photon)的概念.

电子在固态晶体的周期性势垒下能形成电子带隙,类似地,光子晶体中的周期性折射率变化对光产生的布拉格散射可以形成光子带隙.光子带隙是光子晶体最重要的基础特性.而光子带隙所处于的频率范围、频带的宽度以及光子带隙作用的范围,直接关系到光子晶体控制光性能的好坏和能力的大小.本节我们将从多个角度介绍硅基光子晶体带隙方面的知识,并分析当前光子晶体带隙研究的重点.

随着现代微加工技术从微米量级逐渐发展到纳米量级,硅基光子晶体器件中所具有的带隙频率也逐渐增加,其所对应的波长则逐渐减小.从最初的位于中红外 10 μm 波长区,利用大孔硅制作技术,在 1999 年时,逐渐扩展到通信波长区(1.3~1.5 μm)[4],到 2000 年时,更达到了可见光波长区[5].然而,受限于硅基纳米制作技术的成熟度,并考虑到通信波长频率区的重要性与巨大应用前景,硅基光子晶体带隙频率范围的研究目前还是主要集中于通信波长区[6],对于更高频率的硅基光子带隙报道较少.从硅基光子晶体的维度方面来看,硅基光子晶体带隙研究从最初对理想一维与二维光子晶体的带隙理论与实验研究开始;20 世纪 90 年代,发展到全三维硅基光子晶体带隙的理论与实验研究;21 世纪初开始,主

要开始集中于对目前最具实际应用价值的硅基三维结构光子晶体平板的带隙研究[6~10]. 从光偏振角度来看,对硅基光子晶体带隙的研究,从起初对单一横电场带隙或者横磁场带隙的研究;发展到采用各种新方法与原理,追求对所有偏振态的完全光子带隙的研究[11~13]. 从追求硅基光子晶体带隙的光子带隙宽度方面来看,自提出光子带隙概念后,研究人员就一直不断地致力于增大硅基光子晶体的光子带隙宽度[6~10]. 在这方面,各种各样的光子晶体新构型近年来层出不穷,不断涌现,一直是光子带隙研究领域内研究的热点问题[13~19].

通过综合分析发现,在对硅基光子晶体带隙的研究中,目前研究人员主要关心的是硅基平板光子晶体的带宽以及偏振问题. 因此,如何提高硅基光子晶体光子带隙的宽度,以及如何实现完全带隙,是目前研究的核心问题. 相对地,在现有的硅基微纳加工技术条件下,对硅基光子晶体的带隙频率所处范围的研究已经相对比较成熟,因此目前很少再有这方面的研究报导. 所以,本节在描述了光子带隙的基本概念及其产生原理之后,将特别介绍一种新型的硅基绝缘体二氧化硅上硅(SOI)环形光子晶体平板结构. 这种硅基 SOI 环形光子晶体平板与普通硅基圆形空气孔光子晶体平板相比,在很宽范围的空气体积填充比内,都能够获得带隙增强效果,因此,以此环形光子晶体平板为基础,可以设计和制作出各种高品质因子微腔和宽带光子晶体器件,具有很大的应用潜力与前景.

2.3.1 光子晶体的带隙

电子在固态晶体的周期性势垒下能形成电子带隙,光子在周期性折射率变化中体验布拉格散射以及光子带隙. 频率处在光子带隙中的光子被禁止进入光子晶体. 为了方便起见,下面从最简单的一维层状光栅的带隙产生过程开始介绍.

1. 一维层状光栅的带隙

图 2.29 为光在硅和二氧化硅组成的周期层状结构中的传播原理图与带隙图. 其中,硅的折射率 $n_{Si}=3.4$,而二氧化硅的折射率 $n_{silica}=1.45$,如果光在材料中的半个波长正好等于光栅周期,反射光同相位而相长. 在该结构中,光传输行为由两个因素所决定: ① 层与层之间的非涅耳反射效应,其反射率 $r=(n_{Si}-n_{silica})/(n_{Si}+n_{silica})$; ② 多重散射波与反射波之间的相互干涉的布拉格效应.

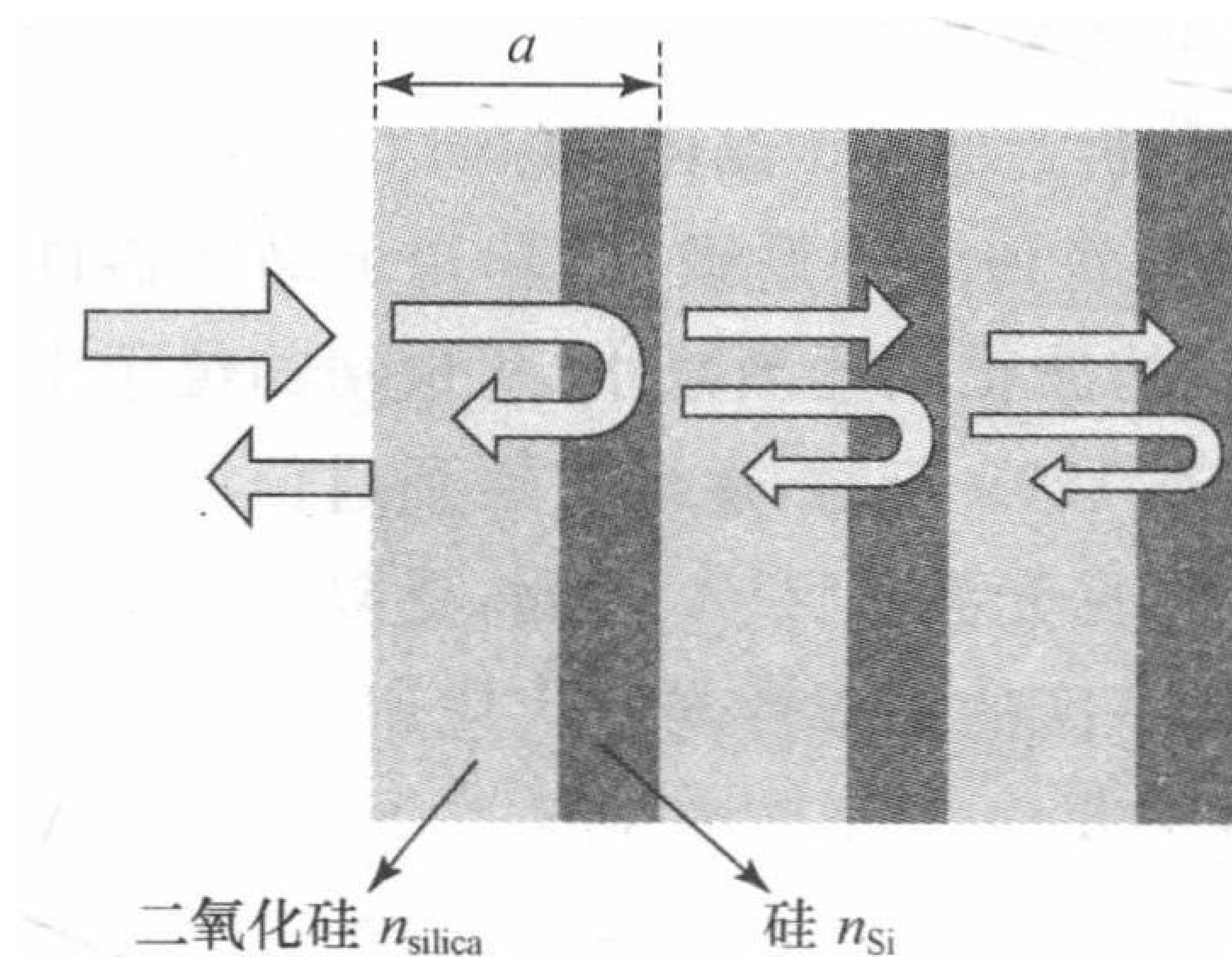


图 2.29 光在硅和二氧化硅组成的周期层状结构中的传播原理图

当波具有波矢 $k=\pi/a$(入射光的波矢和系统的晶格周期 a 相匹配)的时

候，或者说，当周期正好为光在材料中的波长的一半的时候，布拉格效应最强烈.需要注意的是，菲涅耳反射效应是和特定波长没有关系的，而布拉格效应是波长依赖的. 这两种效应联合作用的效果形成了图 2.30 所示的能带结构图：其禁带范围在 $\omega a/(2\pi c)=0.180\,379$ 和 $\omega a/(2\pi c)=0.309\,441$ 之间. 此时，不在禁带频率范围内的光可以在该结构中传输；相反，处于禁带频率范围内的光就会被反射出来. 带图遵循 $c_m=\lambda_m\nu=2\pi\nu\lambda_m/(2\pi)=\omega/k$ 的基本关系式，也就说在某一材料中以速度 c_m 传播的光子在 ω-k 的色散图中会被描绘成一条简单直线. 当在带边的时候，由于存在布拉格效应，这条色散直线会被打断，并且会出现光子禁带. 频率处在光子带隙中的光具有高反射率，因而被禁止进入光栅中. 光子带隙的出现可以用两种模型来描述：菲涅耳反射效应和布拉格效应；光程差模型.

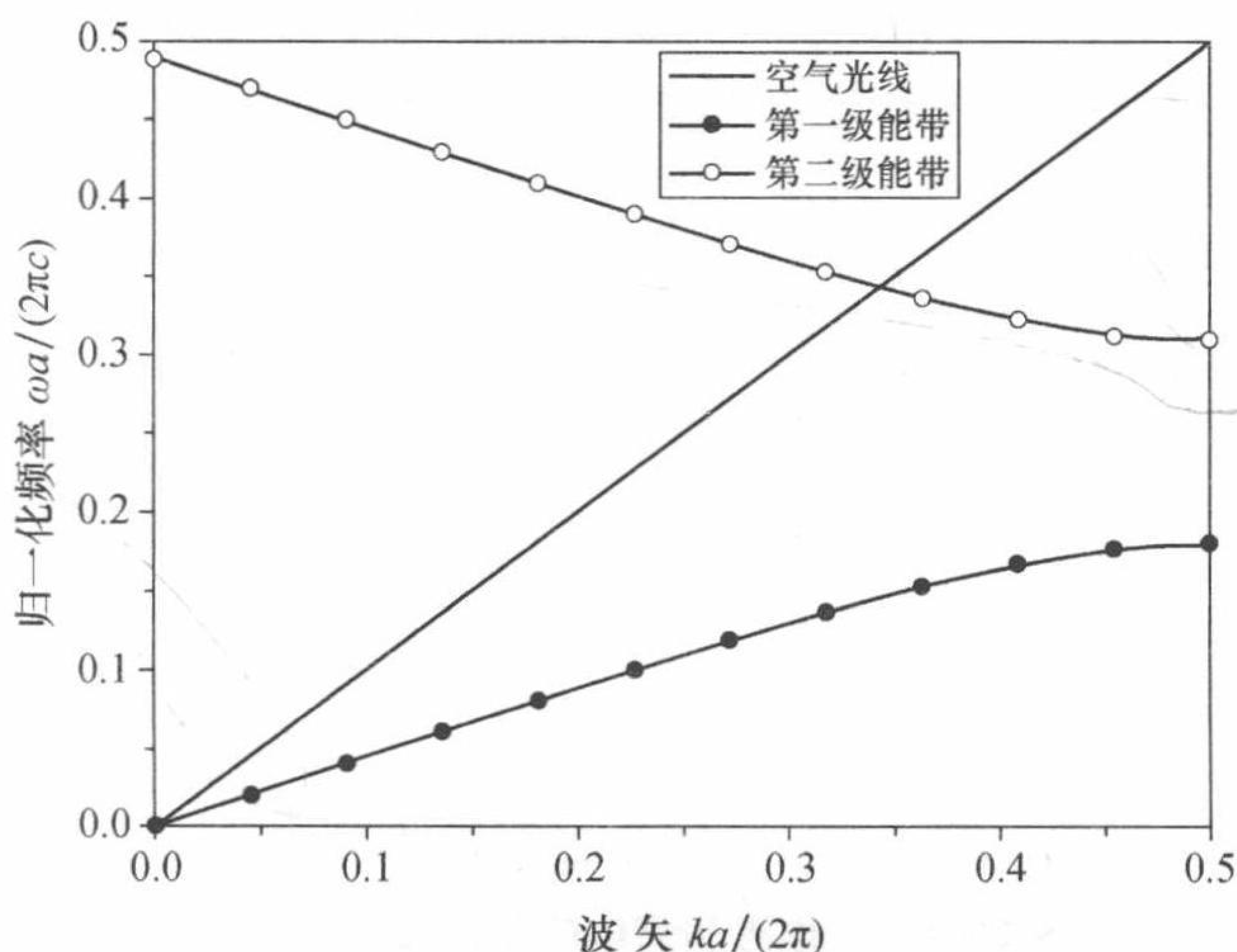


图 2.30 一维层状光栅结构的光子能带图，结构参数和图 2.29 中所示的一样

从菲涅耳反射效应和布拉格效应的角度来看，这种结构中带隙的出现是这两种效应联合作用的结果. 对于低折射率差的周期结构系统，菲涅耳反射很弱，所以要很多次的干涉才能达成较高的反射率. 自然地，这时候有很强的波长依赖性，因此，系统主要体现出与波长相关联的布拉格反射效应，光子带隙的宽度也较窄. 与之相对应，对于高折射率差的系统，菲涅耳反射会很强烈. 因此，入射到系统中的光在仅经历很少几个周期后就能够完全反射. 这时候，系统主要体现出与特定波长无关的菲涅耳反射，并且光子带隙的宽带会较大，即会出现较大的光子带隙.

从光程的角度来看，光模式的有效折射率是由它在结构中的场分布决定的. 分析能带边缘的场分布情况，光场要么存在于低折射率区，要么存在于高折射率区. 因为周期结构的晶格周期 a 和几何长度是确定的，带边波长的不同

必然会伴随着光程差的不同，因此，其有效折射率也会不一样. 实际上，光场的分布和折射率的分布非常相似，特别是当每层的厚度都等于材料中 1/4 波长的时候，$\lambda_l/n_{h,eff}=\lambda_s/n_{l,eff}$，即 $\lambda_l/\lambda_s=n_{h,eff}/n_{l,eff}\approx n_h/n_l$，也即，下带边和上带边波长的比是由折射率的比决定的. 折射率差越大，光子带隙的宽度越大. λ_l 和 λ_s 分别为带边的长波长和带边的短波长. 有效折射率 $n_{h,eff}$，$n_{l,eff}$ 分别由晶格周期内相应模式的场分布情况决定. 需要说明的是，上式中的近似对于每层厚度偏离前述四分之一波长尺寸过大的情况不再适用. 对于偏离较大的情况，也可能会出现光子带隙，但是光子带隙的宽度要远小于折射率的比.

2. 硅基二维光子晶体带隙

现在，我们把上述一维层状周期结构的带隙观点引入到二维周期结构中来. 在二维光子晶体结构中，一维情形中得到的高折射率差带来宽光子带隙宽度的结论，也被扩展到了角度带宽中. 也就是说，对于某一特定波长，在很大的角度范围内，光都会被反射.

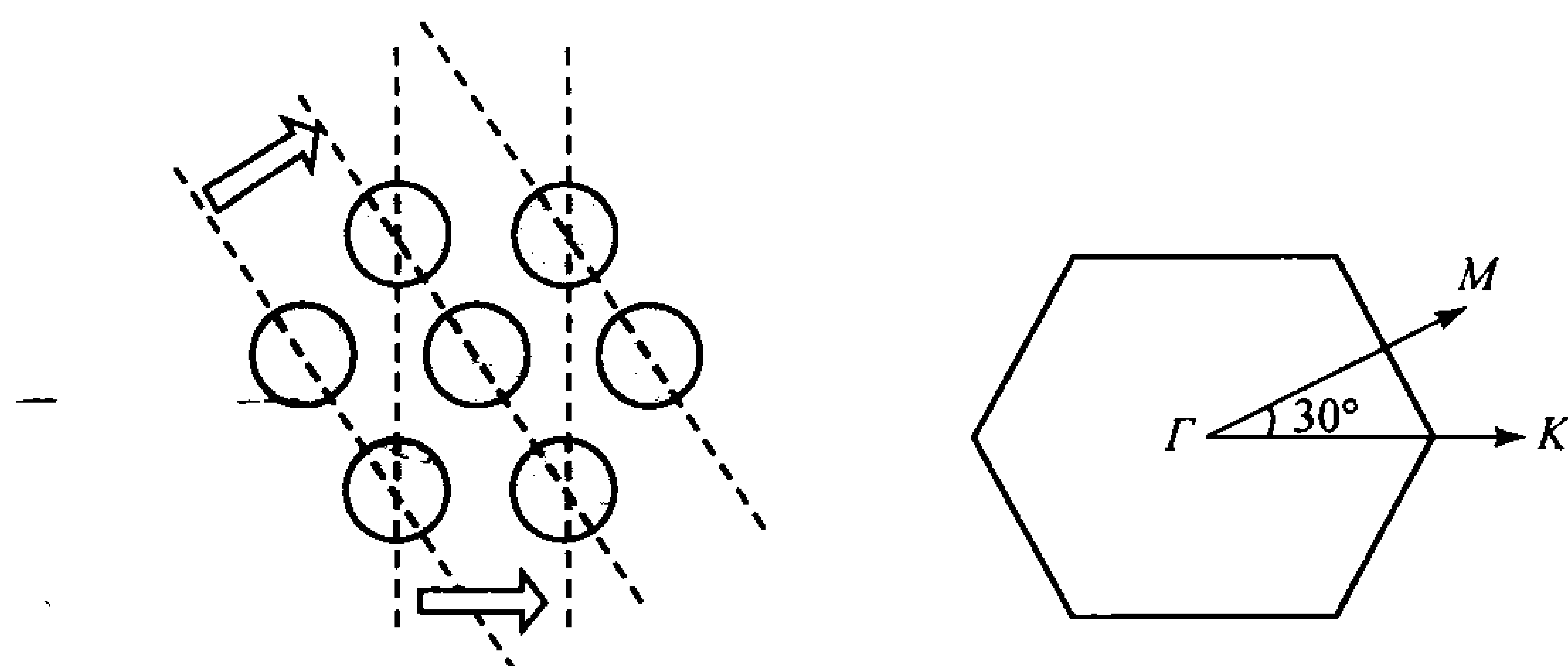


图 2.31 二维光子晶体的晶格和布里渊区

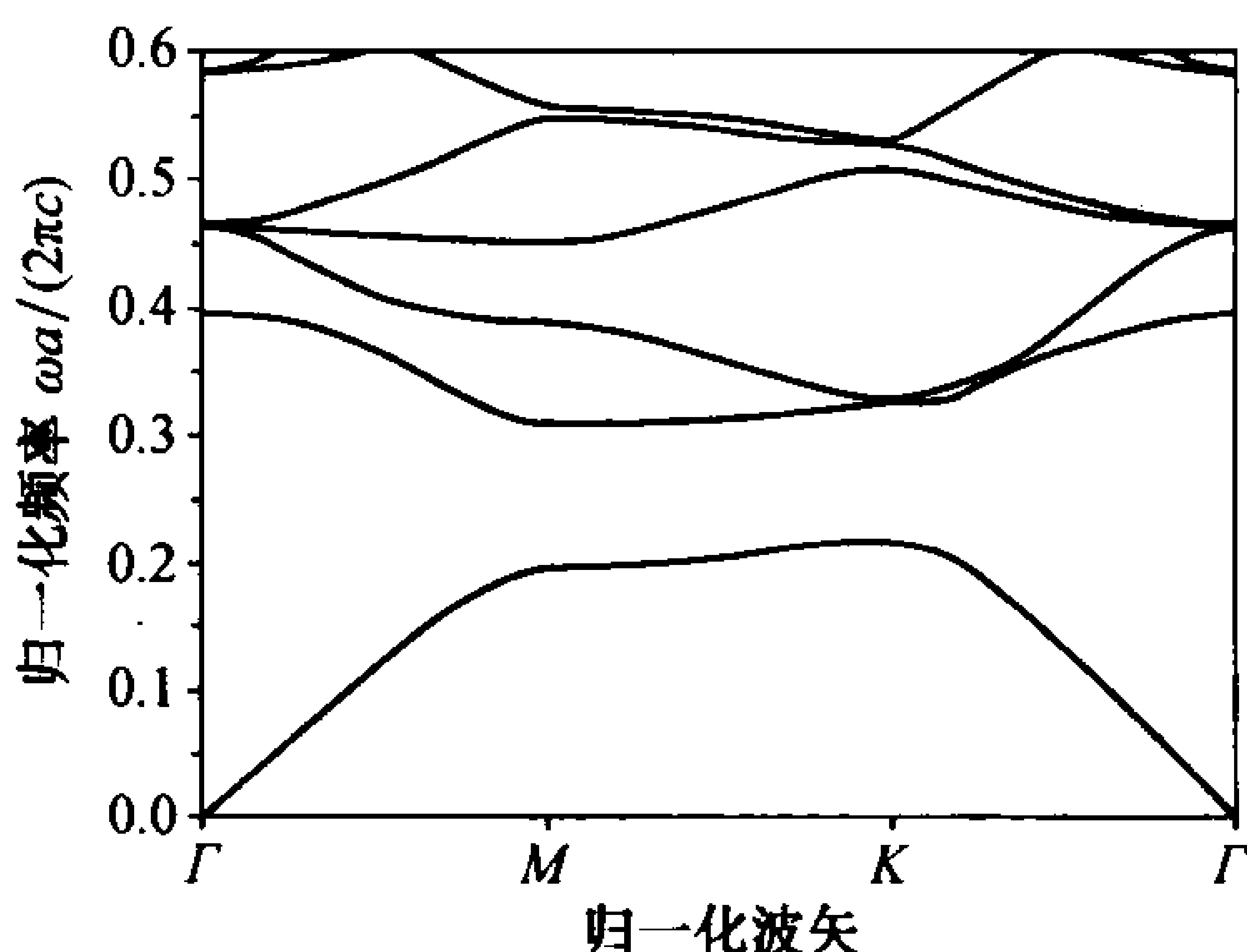


图 2.32 二维光子晶体的横电模(电场主要在光子晶体平面内)能带结构图

因为如图 2.31 所示，二维光子晶体可以看做是由两个或者多个一维光栅结构叠加复合而成的. 当考虑不同的光传播方向时，光经过的实际光栅周期会不一样. 这样，不同周期方向的一维光子禁带会出现在不同的光谱频率位置. 如果光子带隙足够宽，光子带隙光谱频率会出现重叠，这就意味着该频带范围内的光在这几个方向上都被反射. 这个光谱区就被称为二维光子晶体的光子带隙. 如图 2.31 中所示的三角

晶格所示，每个有效周期都在重复着60°的结构；所以如图中所示的Γ-K(0°)和Γ-M(30°)，光传递方向最大的角度差异只有30°. 把不同方向的一维光子带图集中画在一起，就得到了图2.32中所示的二维光子晶体能带结构图. 该能带对应的光子晶体结构参数为，空气孔填充因子40%，空气折射率取为1，背景硅材料的折射率为3.4. 图中阴影区代表着光子带隙区，也即该区域内各个方向(Γ-M和Γ-K)的光子禁带重叠，其禁带范围在$\omega a/(2\pi c)=0.218\,63$和$\omega a/(2\pi c)=0.310\,25$之间.

3. 硅基二维光子晶体平板带隙

进一步地，我们来考虑硅基光子晶体平板的带隙. 与二维光子晶体不同的是，平板结构具有有限的平板高度. 因此光子晶体平板实质上是个三维结构，二维光子晶体的计算只适用于平板高度是无限高的情况，另外，二维光子晶体的带结构只相当于垂直方向的波矢等于零的情况，这一切要求对于有限高的光子晶体平板必须采用三维方法计算[9].

对于光子晶体平板，不仅必须要考虑光子晶体周期平面内的光限制，还要考虑在垂直与平板方向上的光限制. 而垂直方向上的光限制主要依靠平板芯层与上下包覆层之间的折射率差所产生的菲涅耳全反射. 所以，考虑光子晶体平板的时候，在二维光子晶体的带结构图中要叠加反映垂直光限制的色散曲线，通常被称为光锥线. 光线以上的区域称为光锥. 位于光锥中的模式称为辐射模，其在光子晶体中传播的时候，会泄漏到包覆层中去，所以这种模式的损耗很大，一般不用于光器件中.

因此，光子晶体平板的带结构计算分为两个阶段. 第一个阶段计算光子晶体平板中周期元胞中存在的光子态，得到初步能带图，这个阶段考虑了光子晶体平板平面内的周期结构光限制；第二个阶段计算光锥并和第一阶段的能带图重叠，并把光锥的区域标记为不透明区，这一阶段考虑垂直方向折射率差所产生的菲涅耳反射对光的限制，并和第一阶段的结果结合起来. 这样，就得到了光子晶体平板的带结构图. 第一阶段的内容和二维光子晶体的情况类似，所以这里主要介绍对光锥的处理情况. 对于宏观背景来说，光锥就是所有在该背景材料中可能支持与存在的模态，通常只要计算出光锥的下边界就可以了，因为高频模态肯定包含在这个下边界之上. 当背景为均值材料时，下边界可以用k/n表示，即为波矢除以材料的折射率. 当背景为周期结构时，下边界就是该二维周期结构能带图中最低的那条色散曲线. 一个完整的硅基二维光子晶体平板和其所对应的光子能带图如图2.33所示，该空气桥硅基二维光子晶体平板的结构参数为：平板高度$h=0.7a$，空气孔填充因子为35%，空气折射率为1，硅折射率为3.4，其禁带范围在$\omega a/(2\pi c)=0.250\,14$和$\omega a/(2\pi c)=0.338\,82$之间.

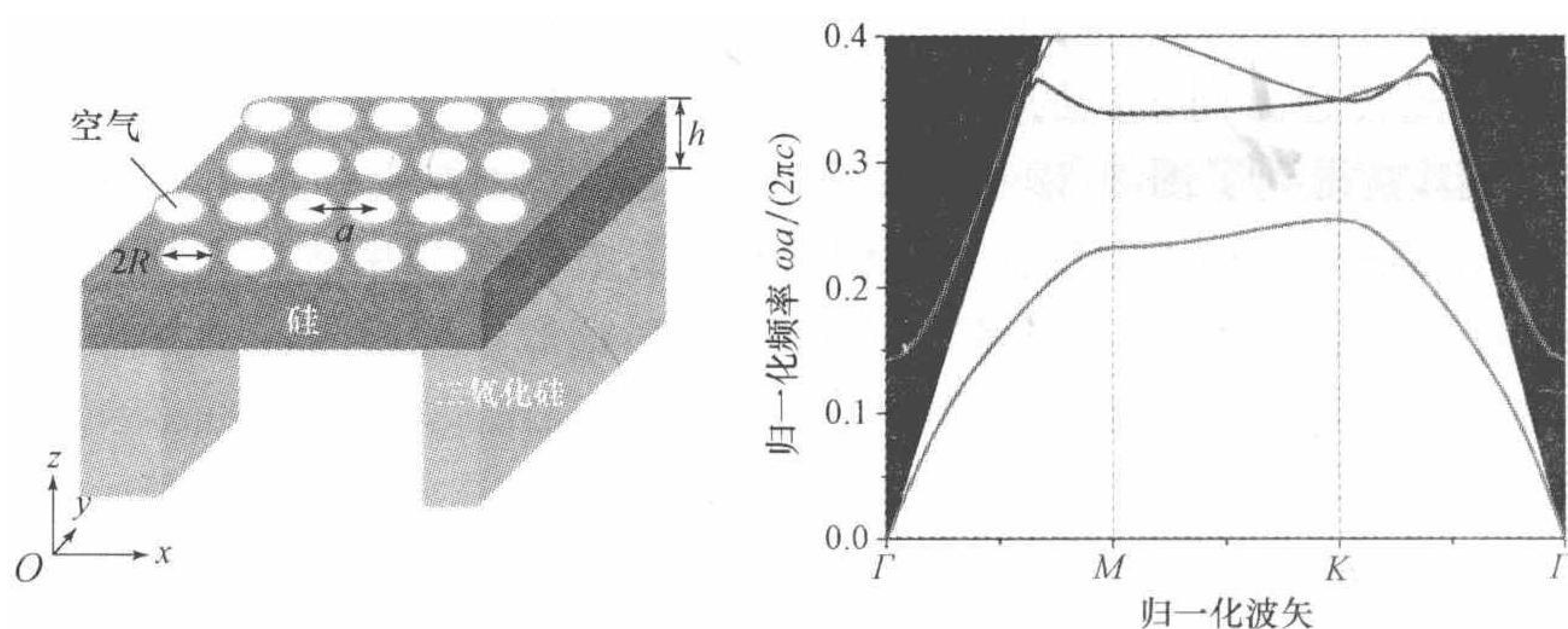


图 2.33 空气桥型硅基光子晶体平板示意图和其横电模(TE)能带结构图

2.3.2 硅基环形光子晶体的带隙增强

光子带隙是光子晶体的核心与基础特性，利用光子晶体的带隙特性可以产生出多种有用的应用[8,11,13,20~27]，例如，把光控制到光子晶体的线缺陷中形成光波导，或者把光控制到光子晶体的点缺陷中形成光学微腔[23~25,28]. 当光子晶体的光子带隙的带宽增加时，这些属性会更加具有意义，因此人们一直致力于增大光子晶体的带隙宽度. 环形光子晶体就是把介电质柱型光子晶体和空气孔型光子晶体结合在一起，因此，这种光子晶体兼有两者的优点并被首先用来实现更宽的完全带隙，即对某个特定的频率范围，对横电(transverse electric, TE)偏振态和横磁(transverse magnetic, TM)偏振态同时具有光子带隙. 通过使用电子束光刻和反应离子束刻蚀[23]，空气孔型环形光子晶体已经被成功制作了出来；而通过原子层沉积和牺牲层刻蚀技术，更加一般的环形光子晶体型态也能被制作出来[29]. 更重要的是，因为环形光子晶体中有两个可以同时调节的结构参数，比通常用的圆形空气孔型光子晶体结构多一个自由度，因此环形光子晶体结构能够很方便地用来调节光子带结构而获得特定的色散特性，如自准直特性和慢光特性[23,27,28].

1. 结构

非对称绝缘体二氧化硅上硅平板是实现硅基光电集成回路最流行和最具有潜力的材料平台之一，所以我们以此为硅基环形光子晶体平板的结构基础来分析. 硅基环形光子晶体平板的具体结构如图 2.34 所示，在折射率为 1.45 的二氧化硅基底上面是硅基三角晶格环形光子晶体平板. 环形空气孔的外环和内环半径分别用 R 和 r 表示，空气的折射率为 1，平板的其余部分都是折射率为 3.4 的半导体硅；光子晶体的晶格常数为 a，平板的厚度为 h. 为了获得好的光限制效果和较大的光子带隙宽度，厚度 h 设为 $0.6a$.

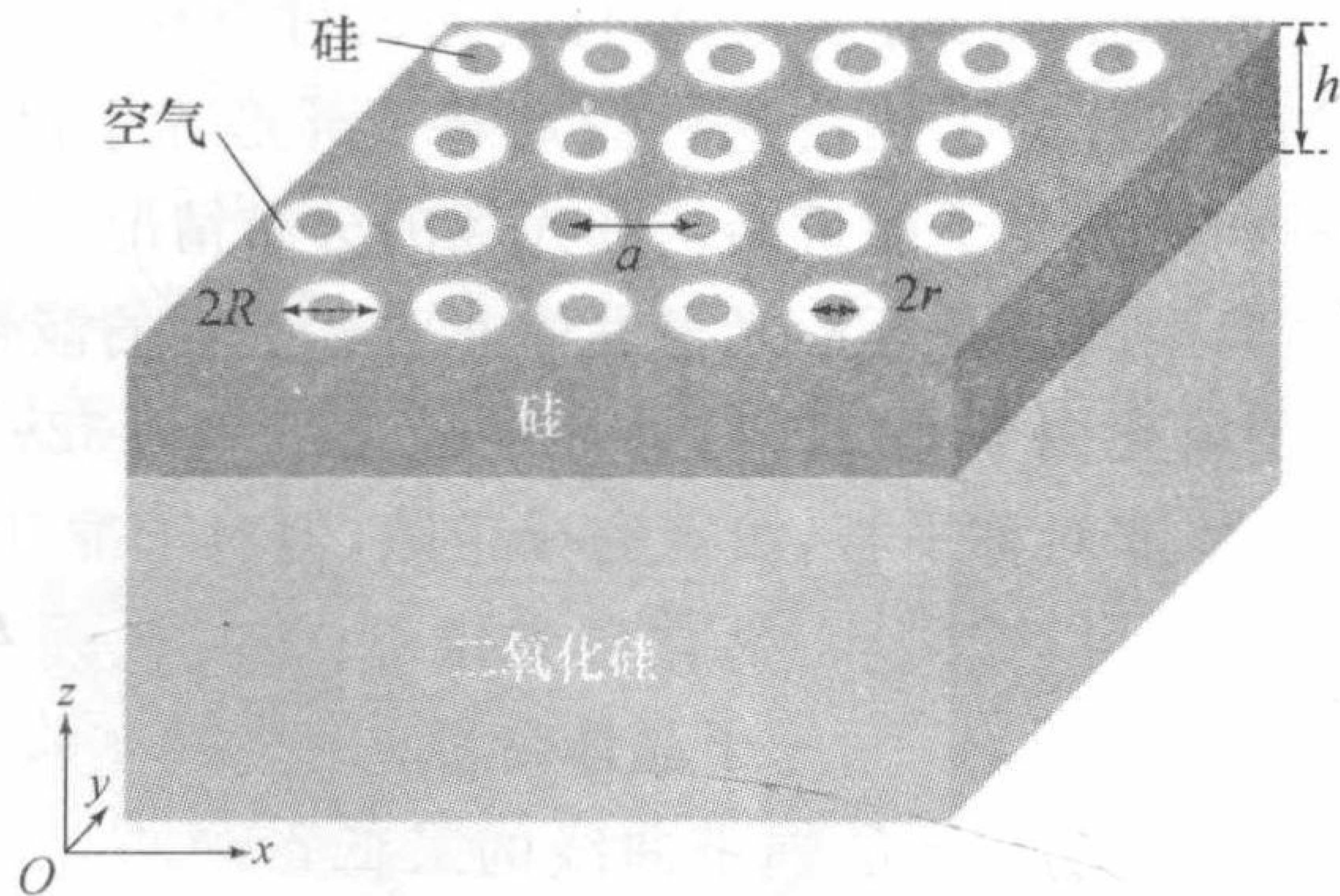


图 2.34　厚度为 h、晶格常数为 a 的硅基三角晶格非对称环形光子晶体平板结构原理图

另外，为了获得高效的计算，二氧化硅包覆层和硅平板上的空气包覆层的厚度都设为 1.5a. 这个包覆层的厚度选取是在对比了更厚的包覆层如 2.0a 和 2.5a 之后做出的. 如图 2.35 所示，在二氧化硅光锥以下，各种不同包覆层厚度设置计算得到的光子能带结构图近似完全一样. 该能带结构图中不同包覆层厚度的光子晶体平板结构的其他参数都一样：光子晶体结构的空气体积填充因子是 20%，外环半径是 $R=0.32a$. 因此，我们采用 1.5a 的包覆层厚度获得的结果是合理的.

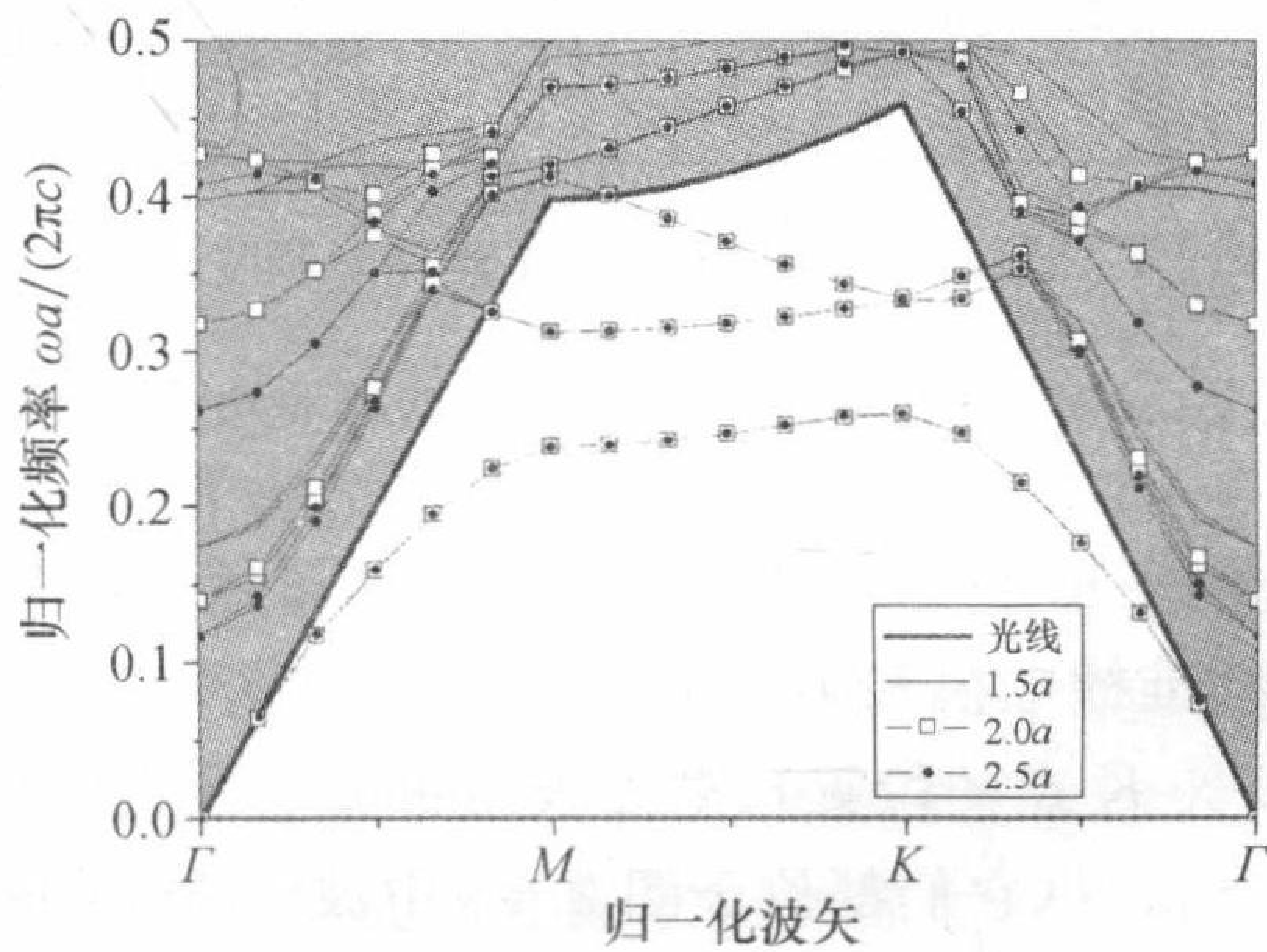


图 2.35　不同包覆层厚度的非对称环形光子晶体平板能带结构比较图

2. 能带分析

对于对称型结构的光子晶体平板，如空气桥式平板，其平板内的光模态可以划分为横电模和横磁模两种偏振态[21]. 然而对于非对称光子晶体平板，由于在垂直光子晶体平面的 z 轴方向上没有镜面对称性，因此横电模和横磁模会存在混杂. 但是习惯上，非对称平板光子晶体的模式仍然分为准 TE 模式和准 TM 模

式[8,22]，因为当平板厚度很薄的时候，这种标记方式可以在 TE-TM 混杂很低的时候判别主要模式的状态. 为了确信这种假定对于环形光子晶体非对称平板也是合适的，首先分析和研究光子能带结构和光场的分布情况.

图 2.36 为一个典型的硅基环形光子晶体非对称平板的能带结构图，该光子晶体结构的空气体积填充因子是 20%，外环半径是 $R=0.32a$. 从图中看到，对于该硅基环形光子晶体非对称平板，没有完全带隙，但对于准 TE 的偏振模式存在光子带隙. 从图 2.36 可以看到，二氧化硅光锥下面的准 TE 带隙由最下面的两条准 TE 色散曲线决定. 具体的，光子晶体带隙的带宽是由第一条准 TE 的能带曲线的最高点和第二条准 TE 的能带曲线的最低点确定的. 在这个带隙中间另外有两条准 TM 能带曲线. 因此，研究这四条能带曲线应该可以很充分的验证光子带隙的特性.

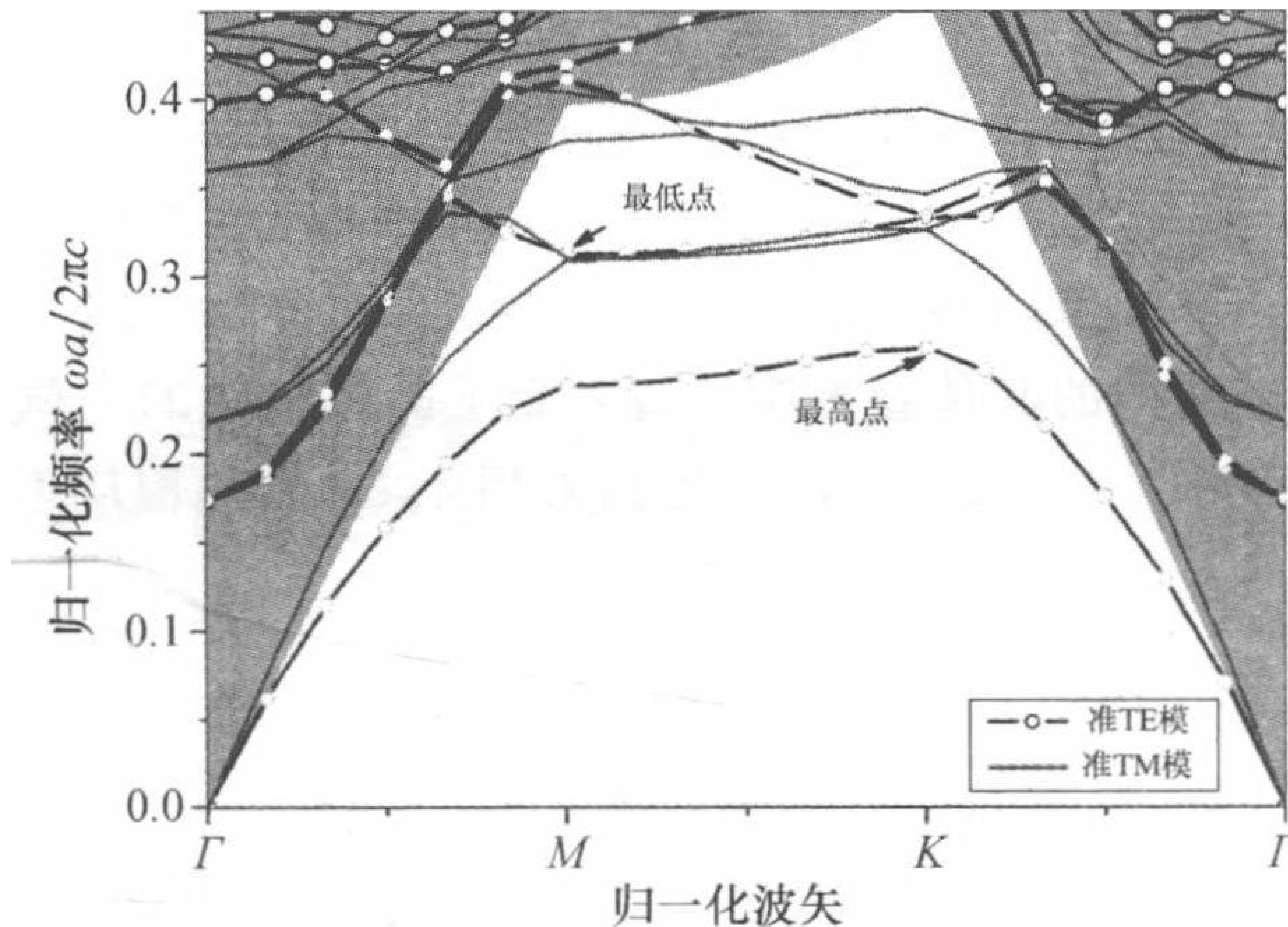


图 2.36　空气填充比为 20%，外圆半径 $R=0.32a$ 的环形光子晶体平板的能带结构图

现在举一个例子说明. 在二氧化硅光锥下面的区域中，如图 2.37 所示，我们选取第一级能带曲线准横电模和第二级能带曲线准横磁模的最高频率点（它们都位于图 2.36 中波矢 K 点的位置）；第三级能带曲线准横磁模和第四级能带曲线准横电模的频率最低点（它们都位于图 2.36 中波矢 M 点的位置）.

图 2.37 显示了这四个模式的垂直光子晶体平板平面方向的磁场 H_z 的场分布情况. 在每一个子图中，子图的左边部分是以环形的中心做剖面而得到的 Oxz 平面上的 H_z 场分布，右边部分是沿着环形孔的轴线上的归一化 H_z 幅度分布. 图 2.37(a)是第一级能带的布里渊区 K 点波矢的模式分布；图 2.37(b)是第四级能带的布里渊区 M 点波矢的模式分布；图 2.37(c)是第二级能带的布里渊区 K 点波矢的模式分布；图 2.37(d)是第三级能带的布里渊区 M 点波矢的模式分布. 能带级数按照 M 点的频率高低由下到上算起. 图 2.37(a)和(b)中的

H_z 场分布在 z 方向上表现完全对称的分布，说明第一级能带和第四级能带基本上全是偶模，也即是准 TE 模式；相对应地；图 2.37(c)和(d)表现出奇 H_z 分布，说明第二级能带和第三级能带主要是准 TM 模式. 因此，我们可以认为在该结构中，TE-TM 混杂比较轻微，可以基本忽略. 我们也检查了具有其他结构参数和布里渊区内其他波矢位置的模式的场分布情况，和前述状态均表现一致的特性.

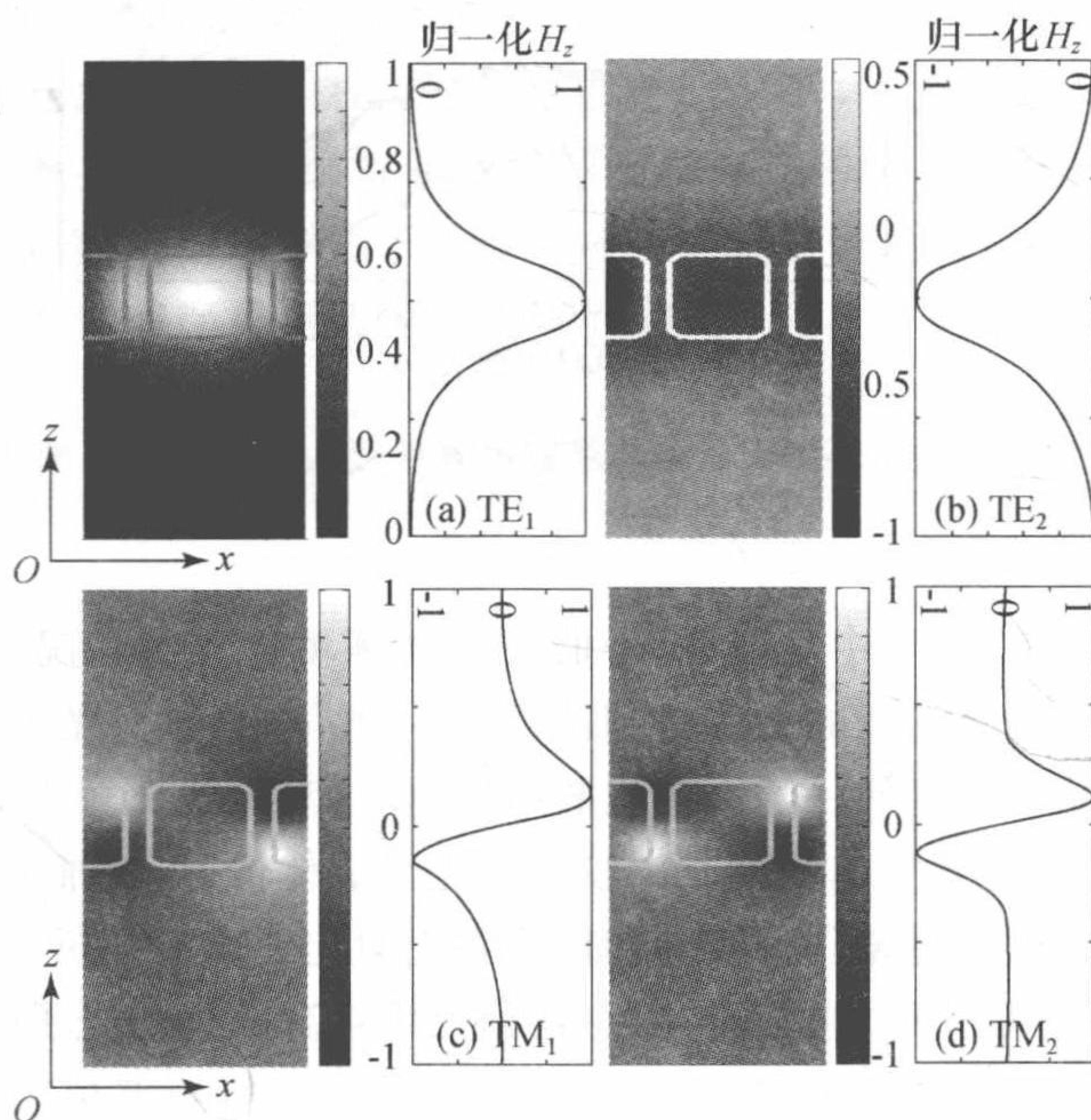


图 2.37　准 TE 模和准 TM 模在一个单元格内的 Oxz 平面 H_z 场分布图

3. 性能分析

在确信了 TE 和 TM 模式的划分在对于非对称环形光子晶体平板结构也是适用的之后，我们研究了环形光子晶体的光子带隙随着结构参数变化而改变的规律情况. 图 2.38 显示了对于空气填充比在 15%～40%范围内以 5%为增量变化的一系列硅基环形光子晶体非对称平板，其光子带隙归一化带宽(带隙宽度与带隙中间频率之比)随着环形孔的外圆半径 R 变化而改变的关系图. 图 2.38 中的标记 f_{3D}，f_{2DE} 和 f_{2D} 分别代表用全三维方法计算的硅基二维环形光子晶体非对称平板的空气体积填充比、用二维有效折射率方法计算的硅基二维环形光子晶体非对称平板的空气体积填充比和用二维方法计算的硅基二维环形光子晶体的空气体积填充比. 图中，每条曲线最左边的一点对应于通常圆形空气孔光子晶体(内圆半径为 0)的状态. 用全三维平面波展开方法计算的空气体积填充比的最小值(15%)和最大值(40%)所得的光子带隙和用二维有效折射率法以及用二维平面波展开法获得的结果也进行了比较，如图 2.38 中所示.

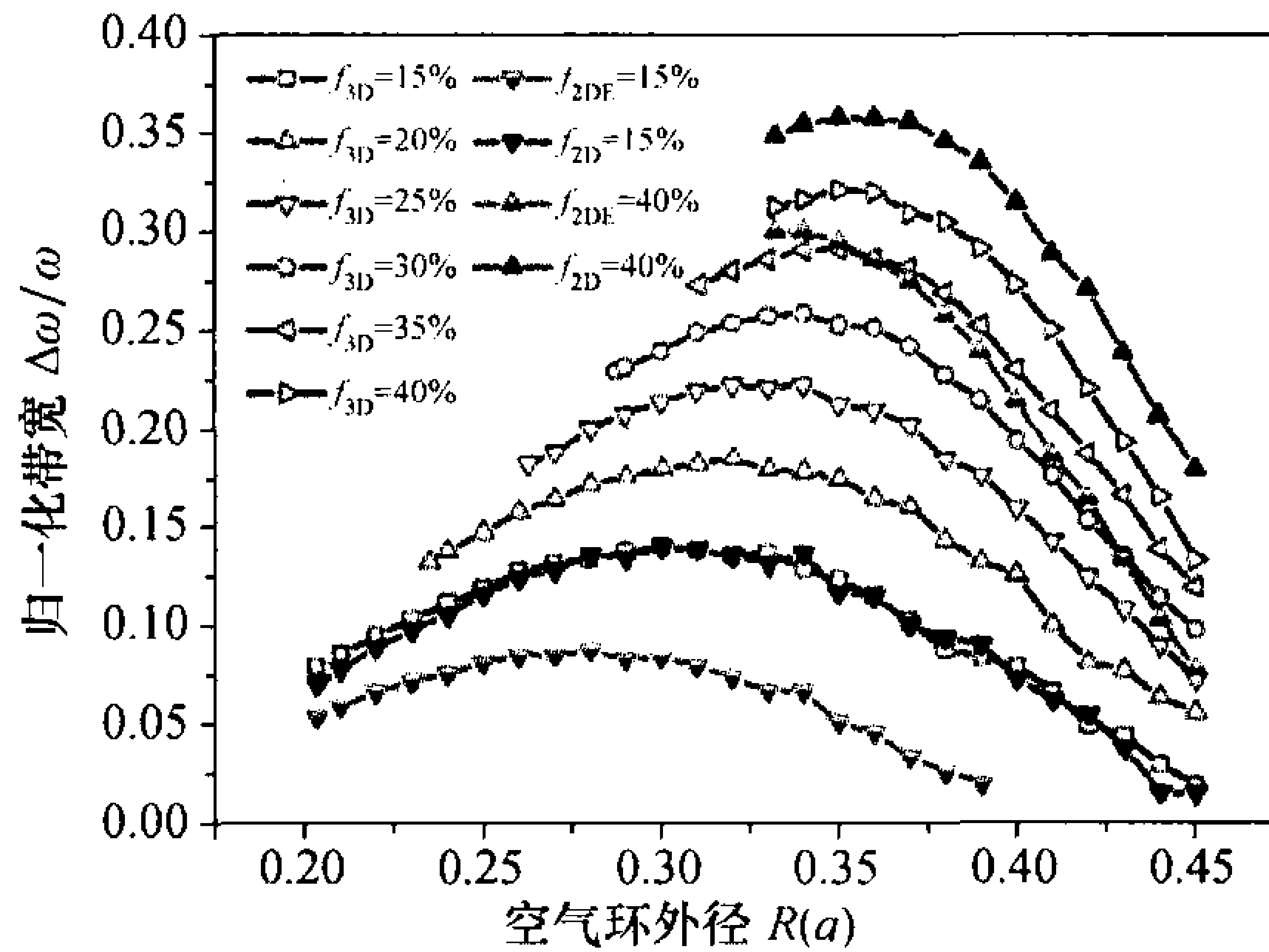


图 2.38 归一化硅基环形光子晶体带隙随着环形孔外圆半径 R 变化图，单位 a 为晶格常数

我们可以从图 2.38 中发现，对于低空气体积填充比的情况，用二维方法计算的结果几乎和用全三维的方法得到的结果一样. 这是因为光子晶体平板结构只是二维光子晶体结果的一种小微扰的结构，在这种微扰比较小的情况下，结果影响不大. 但是对于空气体积填充比较高的光子晶体情况，二维方法比三维方法的结果得到的带隙要更宽广一些. 同时，当我们对比二维有效折射率方法的结果时，二维有效折射率获得的光子带隙总是比三维情况获得的光子带隙要窄. 这是因为在二维有效折射率方法中，一般是使用一个有效折射率值来等效光子晶体平板的背景折射率. 而通常，这个有效折射率都会低于背景折射率，所以获得的能带曲线一般较低. 等效折射率方法只能获得计算有效折射率频率附近的近似能带结构，对于大带宽的情况，其结果就不再那么准确[30]. 这也是我们要在这里使用全三维方法研究硅基环形光子晶体非对称平板的原因，同时这也解释了图 2.38 中出现的主要现象. 也就是说，二维方法可以便捷地定性研究光子晶体，便于人们理解光子晶体的特性；与之相对应地，对于光子晶体带隙的定量分析，三维方法是必须的.

对于某一特定的空气体积填充比，归一化硅基环形光子晶体带隙随着环形孔外圆半径 R 变化曲线展现出单峰曲线，这表明通过在大范围内调节环形光子晶体的参数，光子晶体在光子带隙附近的能带结构没有随着 R 的变化而存在性质上的改变，并且能够找到光子带隙的最大值. 这就是说，对于某一特定空气体积填充比，采用环形空气孔光子晶体能够提高常规圆形空气孔光子晶体的光子带隙. 如图 2.39 所示，同曲线最左边的那点所代表的常规圆形空气孔光子晶体相比，虽然对于高空气体积填充比 40%的情况，环形空气孔光子晶体只有 1%的带宽增加；但是值得注意的是，对于低空气体积填充比 15%的情况，常规硅基圆

形空气孔光子晶体平板能够获得 7.8%的光子带隙宽度，与之相比，优化的硅基环形空气孔光子晶体平板能够提高光子带隙宽度到 14.0%，调高了接近一倍.因此，环形光子晶体对于既要求要有较大的光子带隙，又要求有较低空气体积填充比的应用(如光学微腔等)特别有用.对于常规硅基圆形光子晶体，如果空气体积填充比较大的话，光子晶体平板垂直方向的光限制就会变得较弱，这样就会限制光学微腔最大可能获得的品质因子 Q 值.因此，通常在最大光子带隙和最小空气体积填充比之间要互相权衡[31].在这种情况下，环形光子晶体的特性就表现出了非凡的意义.它能在维持低空气体积填充比的情况下增强光子带隙的宽度；或者也可以这样说，环形光子晶体平板在提供常规圆形空气孔光子晶体同样宽度带隙的同时能够保持更低的空气体积填充比.

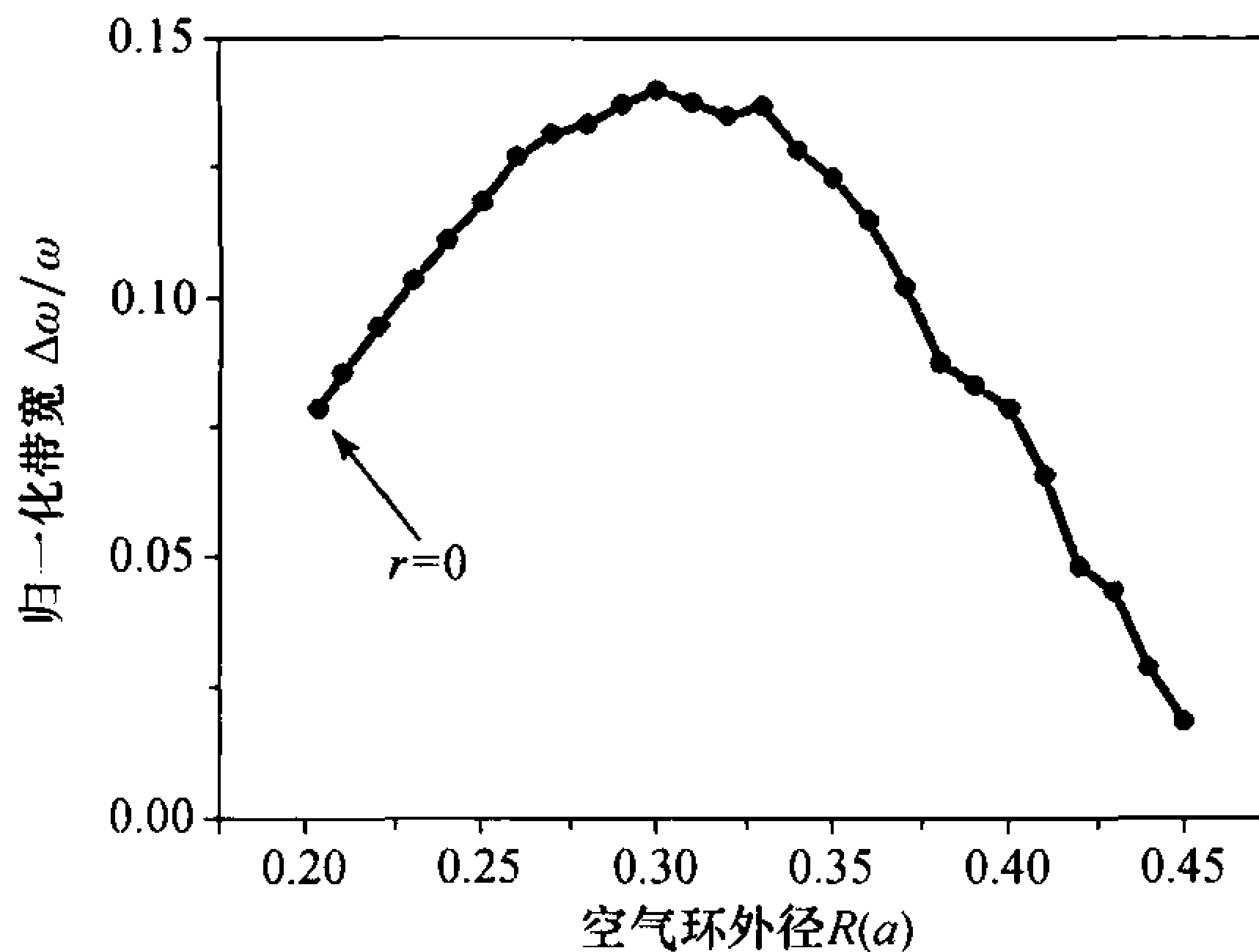


图 2.39　空气体积填充比为 15%时，硅基环形光子晶体平板归一化带隙随着外圆半径 R 变化图，单位 a 为晶格常数

2.3.3　小结

光子带隙是光子晶体最重要的特性.本节在简要介绍了光子带隙产生的机理之后，重点阐述了在硅基光子晶体带隙方面的研究方向与成果，提出了利用硅基环形光子晶体平板来增强常规硅基圆形光子晶体平板带隙的方法；采用全三维平面波展开法，分析了硅基有限高非对称环形光子晶体平板的光子带结构特性；发现在硅基环形光子晶体中，当选取合适的参数时，与具有同等空气填充比的常规圆形空气孔光子晶体相比，能够很大程度上增强光子带隙带宽.对于低空气体积填充比 15%的情况，常规硅基圆形空气孔光子晶体平板能够获得 7.8%的光子带隙宽度，与之相比，优化的硅基环形空气孔光子晶体平板能够提高光子带隙宽度到 14.0%——调高了接近一倍.环形光子晶体的这种特性对于宽带波导、反射镜、传感器、光学微腔等表现出了非凡的意义.

§2.4 硅中光子与载流子的相互作用[1]

半导体受热或光的影响,可以使能量在电学、光学、声学等形式中进行转换.从微观的角度看,这种能量转换是光子通过载流子(电子、空穴)在能级之间的跃迁来进行的.其间,可以有,也可以没有声或其他的中间过程.在宏观上,光子与载流子之间的能量交换则可以用光子的吸收和发射,材料的导电性及折射率的改变来描述.本节将着重介绍光子与载流子之间的相互作用及其对半导体材料及器件性能的影响.

如图 2.40 所示,在半导体中,光子与载流子之间的相互作用可以有如下几种重要的机制.

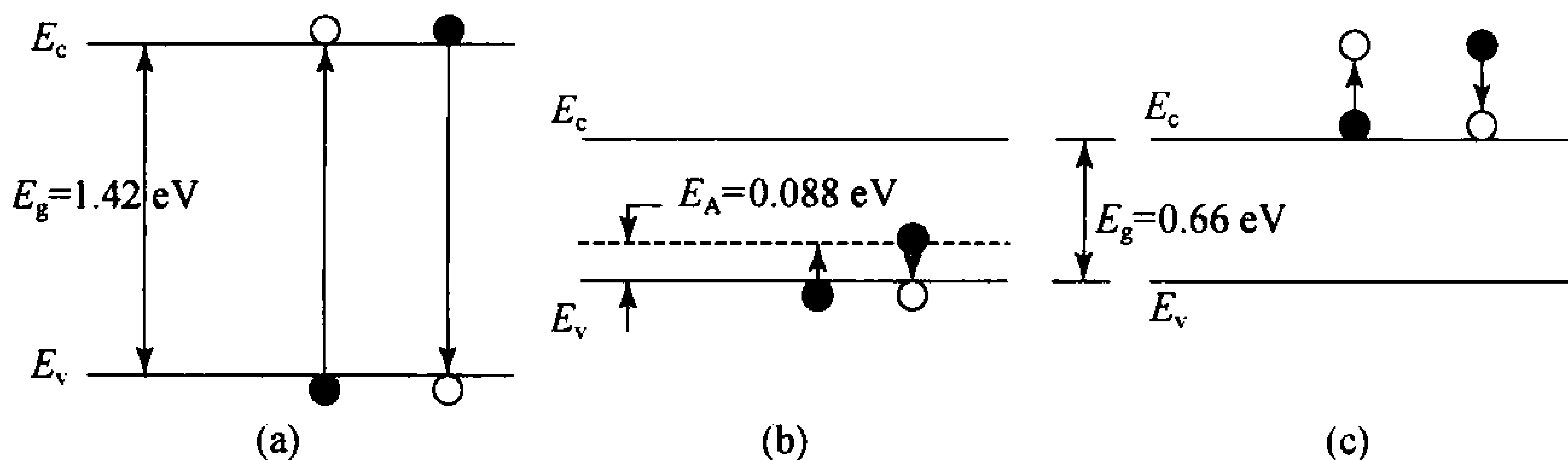


图 2.40 半导体中光子与载流子之间的相互作用

1. 带间跃迁

一个被吸收的光子可以导致在价带中的电子向上跃迁到导带,从而形成一个电子空穴对.电子和空穴的重新复合可以导致发射一个光子.如图 2.40 (a)所示,GaAs 中载流子的带间跃迁可以导致吸收或者发射波长小于 $\lambda_g = hc_0/E_g = 0.87\ \mu m$ 的光子.带间跃迁可能由一个或者多个声子辅助.一个声子是晶格振动的量子,这种晶格振动来自于物体里原子的热振动.

2. 杂质能级与导带(或价带)之间的跃迁

一个被吸收的光子可以导致在施主(或受主)能级与导带(或价带)之间的跃迁.如在一个 p 型材料中,一个低能量的光子可以把电子从导带提到受主能级,在那里电子被受主原子俘获.如图 2.40(b)所示,在 Hg 掺杂的 Ge(Ge:Hg)中,对于波长满足 $\lambda_A = hc_0/E_A = 14\ \mu m$ 的光子的吸收可以导致载流子从价带到施主能级的跃迁.导带中产生一个空穴而受主原子被电离.或者一个空穴可能被一个电离的受主原子俘获;结果就是电子从受主能级退回导带从而与空穴重新复合.这个过程产生的能量可以以辐射的形式发射(以发射光子的形式),也可以以无辐射的方式(声子参与的过程).而跃迁也可以由陷阱态或者缺陷来辅助完成,如图 2.18 所示.

3. 带内自由载流子的跃迁

一个被吸收的光子可以把它的能量传递给给定的某带中的电子，导致电子在该带中跃迁到更高能量的位置. 比如一个导带中的电子可以吸收一个光子以后跃迁到该导带中更高的能级，如图 2.40(c)所示. 接下来则会发生“热化”过程，即电子把能量发射出来并退回到导带底，这部分能量转化为晶格振动能.

4. 声子跃迁

长波光子可以通过直接激发晶格振动的方式来释放它的能量，即通过产生声子的方式.

5. 激子跃迁

吸收一个光子可以导致一个相隔一定距离的电子空穴对，它们之间通过库仑相互作用束缚在一起. 这种独立存在的状态，很像一个氢原子中质子换成空穴的情形，就叫做激子. 当电子和空穴复合时，就会发射一个光子，这就是激子淬灭.

以上这些跃迁都会对半导体材料的总吸收系数产生影响. 图 2.41 展示了 Si 和 GaAs 的吸收系数与光子能量的关系[1]. Si 和 GaAs 的带隙宽度 E_g 分别是 1.11 eV 和 1.42 eV. Si 在 1.1～12 μm 的波段是相对比较透明的，而本征的 GaAs 在 0.87～12 μm 的波段是相对比较透明的.

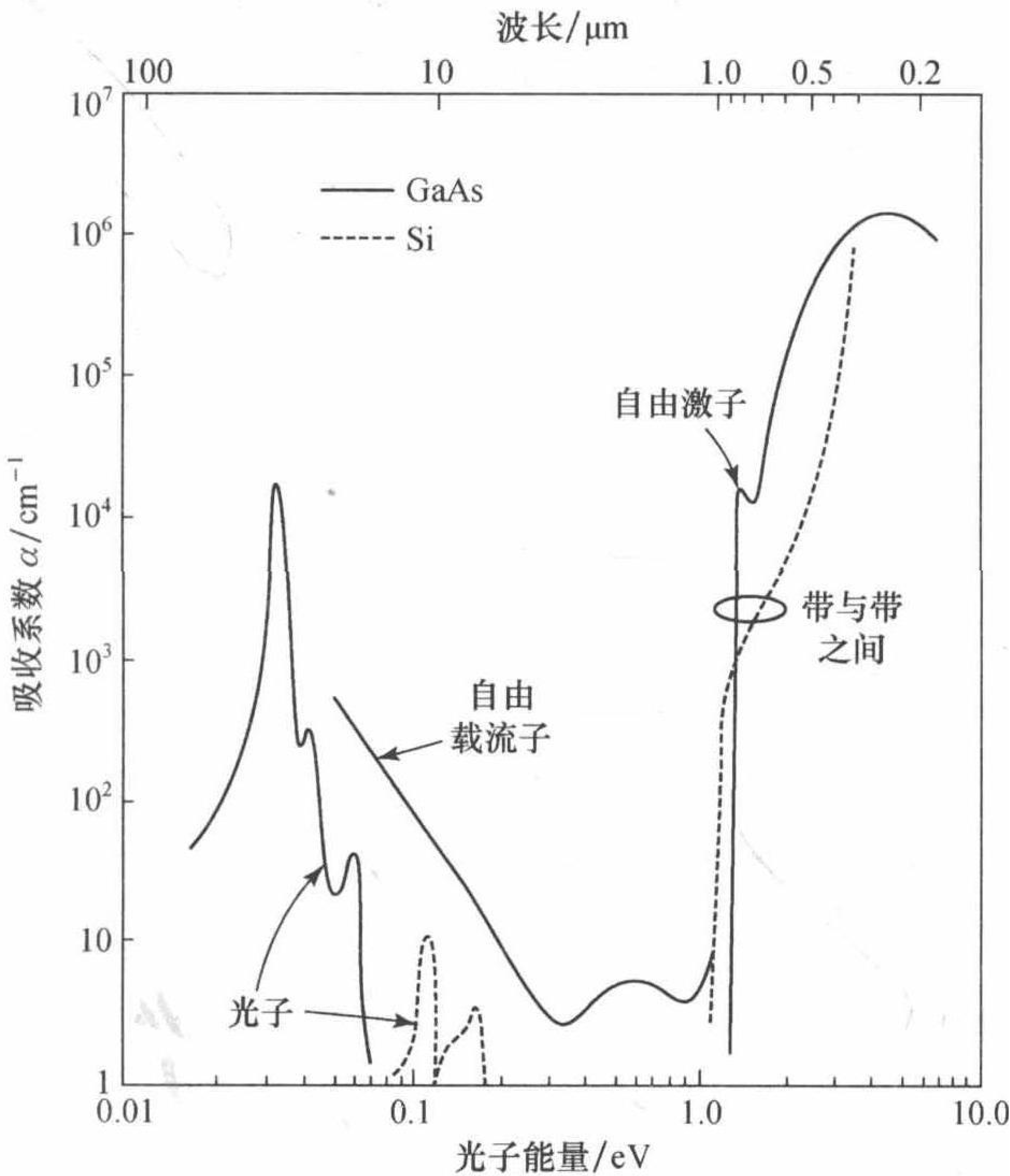


图 2.41　在 $T=300$ K 下，对于 Si 和 GaAs 观察到的吸收系数 α 与光子能量的关系曲线

一般来说，对于能量大于带隙的光子，吸收主要由带间跃迁主导，这是大多数光电子器件的基础．从材料为相对透明的波段($h\nu<E_g$)变化到强吸收波段($h\nu>E_g$)的区域被称为吸收边．直接带隙半导体相对间接带隙半导体有一个更为陡峭的吸收边．下面分三个小节着重介绍载流子跃迁与光子吸收和发射之间的关系．

2.4.1 带间吸收和发射

在忽略其他跃迁形式的情况下，下面介绍有关光子的直接带间吸收和发射的简单理论．

1. 带隙波长

直接带间吸收或发射只发生在光子的频率满足 $h\nu>E_g$ 的情况下．其中最小的频率 ν 即为 $\nu_g=E_g/h$，因此相应的最大的波长为 $\lambda_g=C_0/\nu_g=hC_0/E_g$．如果带隙对应的能量以 eV 为单位(而不是 J)，则以 μm 为单位的带隙波长 $\lambda_g=hc_0/eE_g$ 可以表示成

$$\lambda_g=\frac{1.24\ \mathrm{eV}\cdot\mu\mathrm{m}}{E_g},\tag{2.4.1}$$

其中 λ_g 被称为带隙波长(或者截止波长)．带隙波长 λ_g 可以通过调整由Ⅲ-Ⅴ族元素组成的三元或四元化合物半导体的组分而在相当大的范围内调节(从红外到可见波段)，如图 2.42 所示．

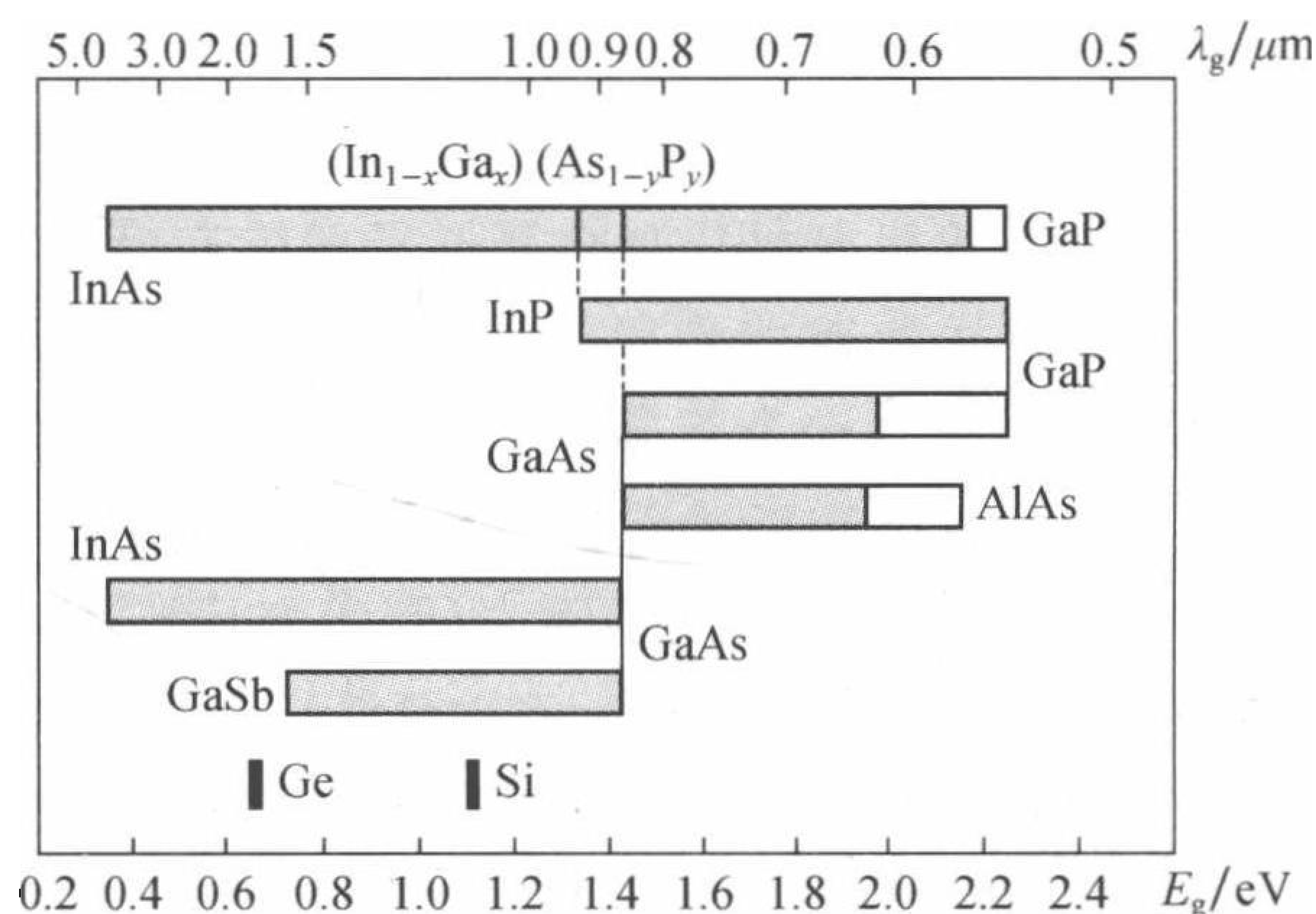


图 2.42 对应于所选择的二元、三元和四元Ⅲ-Ⅴ族半导体的带隙能量 E_g 以及相对应的带隙波长 λ_g，阴影区域表示相对应的材料是直接带隙半导体

2. 吸收和发射

电子从价带激发至导带可以是因为吸收了一个具有适当能量的光子($h\nu>E_g$)，从而产生了电子空穴对，如图 2.43(a)所示．这增加了载流子的浓度

从而增加了材料的电导率. 材料表现得像是一个附加电导和光子流成正比的光子导体,这种效应常常被用于光探测器.

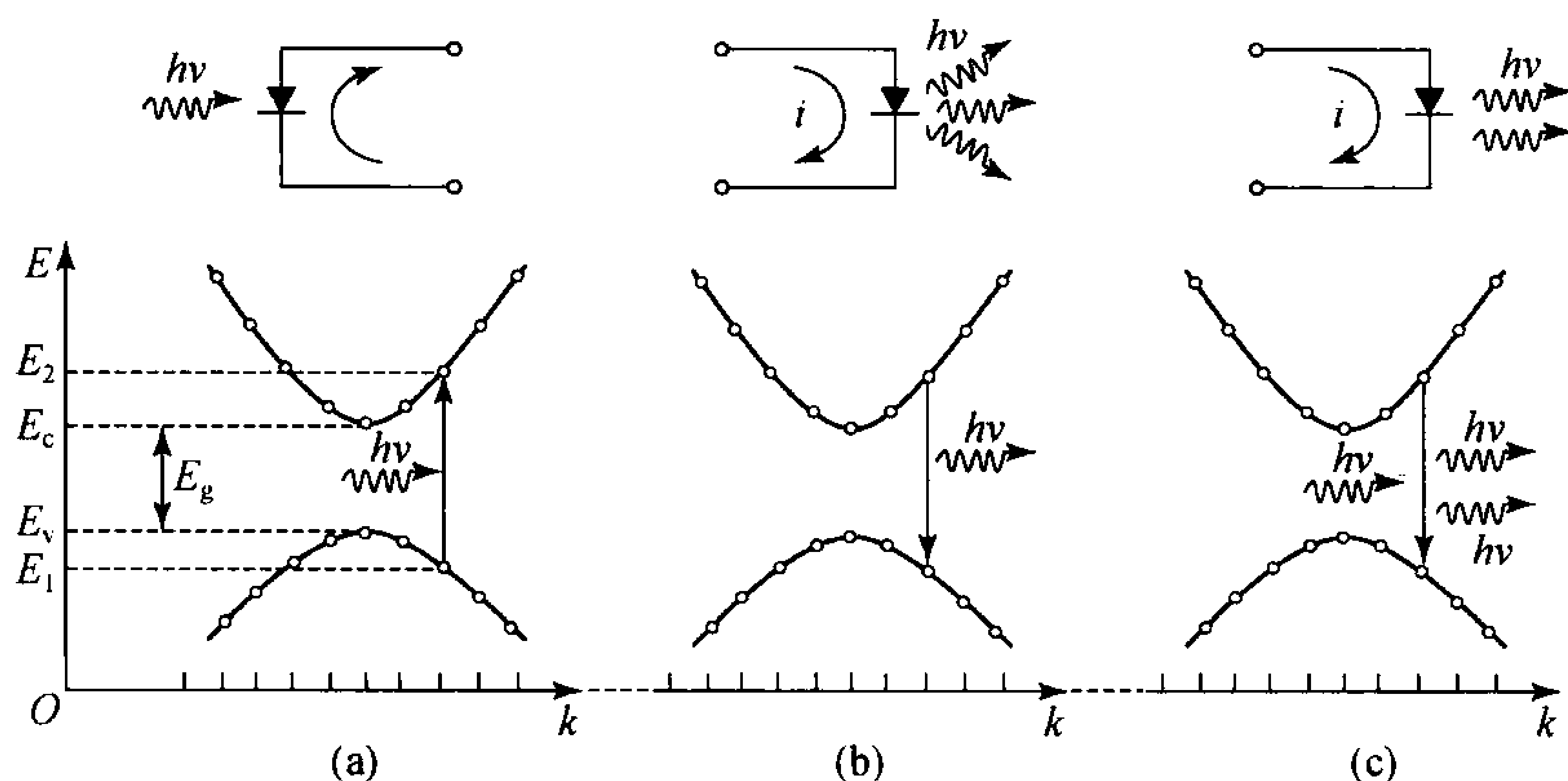


图 2.43　吸收一个光子产生一组电子空穴对(a)；电子空穴对的重新复合导致了光子的自发发射(b)；电子空穴对的重新复合可以由光子激发(c)

电子从导带向价带的退激发(电子、空穴复合)可以导致能量为 $h\nu \geq E_g$ 的光子的自发发射,如图 2.43(b)所示. 如果一个能量为 $h\nu > E_g$ 的光子进入半导体,如图 2.43(c)所示,也可以导致受激的光子发射. 自发的光子发射作为一种重要现象是 LED 工作的基础,而受激的光子发射是半导体光放大器和激光工作的基础.

3. 吸收和发射的条件

(1) 能量守恒. 吸收或者发射一个能量为 $h\nu$ 的光子需要参与这个相互作用两个态的能量(价带的 E_1 和导带的 E_2)差为 $h\nu$. 例如对于电子空穴复合导致的光子发射而言,一个占据 E_2 能级的电子必须和一个占据 E_1 能级的空穴发生作用,这样能量才能守恒,即

$$E_2 - E_1 = h\nu. \tag{2.4.2}$$

(2) 动量守恒. 在吸收和发射光子的过程中动量也必须是守恒的,即 $p_2 - p_1 = h\nu/c = h/\lambda$,或者有 $k_2 - k_1 = 2\pi/\lambda$. 然而,光子动量的量级 h/λ 相比电子和空穴的动量范围来说是很小的. 半导体的 E-k 关系图延伸到 k 为 $2\pi/a$ 的量级,其中晶格常数 a 比波长 λ 小很多,因此 $2\pi/\lambda \ll 2\pi/a$. 这样参与与光子相互作用的电子和空穴的动量大体上是相同的. 这个条件 $k_2 \approx k_1$,被称为 k 选择定则. 满足这个定则的跃迁用垂直的线画在 E-k 图上(图 2.43),表示 k 的变化在 E-k 图的尺度上是可以忽略的.

(3) 与光子相互作用的电子和空穴的动量及能量. 从图 2.43 可以明显地看出,能量和动量的守恒需要一个频率为 ν 的光子去和有着特定能量及动量(由半

导体的 E-k 关系决定)的电子和空穴相互作用. 用式(2.2.1)和(2.2.2)所描述的两个抛物线去近似直接带隙半导体,并记 $E_c-E_v=E_g$,则式(2.4.2)可以写成

$$E_2-E_1=\frac{\hbar^2k^2}{2m_v}+E_g+\frac{\hbar^2k^2}{2m_c}=h\nu, \tag{2.4.3}$$

其中

$$k^2=\frac{2m_r}{\hbar^2}(h\nu-E_g), \tag{2.4.4}$$

$$\frac{1}{m_r}=\frac{1}{m_v}+\frac{1}{m_c}, \tag{2.4.5}$$

$$E_c-E_v=E_g.$$

将式(2.4.4)代入式(2.2.1),则与光子相互作用的能级 E_1 和 E_2 可写为

$$E_2=E_c+\frac{m_r}{m_c}(h\nu-E_g), \tag{2.4.6}$$

$$E_1=E_v-\frac{m_r}{m_v}(h\nu-E_g)=E_2-h\nu. \tag{2.4.7}$$

在特殊情况下,如果 $m_c=m_v$,我们得到 $E_2=E_c+\frac{1}{2}(h\nu-E_g)$,这也是对称性要求的结果.

(4) 光学关联态密度. 它是在直接带隙半导体中,一个能量为 $h\nu$ 的光子在能量动量守恒的条件下参与有关作用的态密度 $\rho(\nu)$. 这个数值包括导带和价带的态密度因此被称为光学关联态密度. 在式(2.4.6)中 E_2 和 ν 的一一对应关系很容易通过关系式 $\rho(\nu)=(\mathrm{d}E_2/\mathrm{d}\nu)\rho_c(E_2)$ 将 $\rho(\nu)$ 同导带中的态密度 $\rho_c(E_2)$ 联系起来:

$$\rho(\nu)=\frac{hm_r}{m_c}\rho_c(E_2). \tag{2.4.8}$$

利用式(2.2.4)和(2.4.6),我们最终得到单位体积单位频率的态密度(光学关联态密度)

$$\rho(\nu)=\frac{(2m_r)^{3/2}}{\pi\hbar^2}(h\nu-E_g)^{1/2},\quad h\nu\geqslant E_g, \tag{2.4.9}$$

如图 2.44 所示,态密度随着 $h\nu-E_g$ 的增加以平方根的关系而增加. 利用式(2.4.7)中 E_1 和 ν 的一一对应关系,以及式(2.2.5)所表达的 $\rho_v(E_1)$,同样可以得到和式(2.4.9)同样的 $\rho(\nu)$ 表达式.

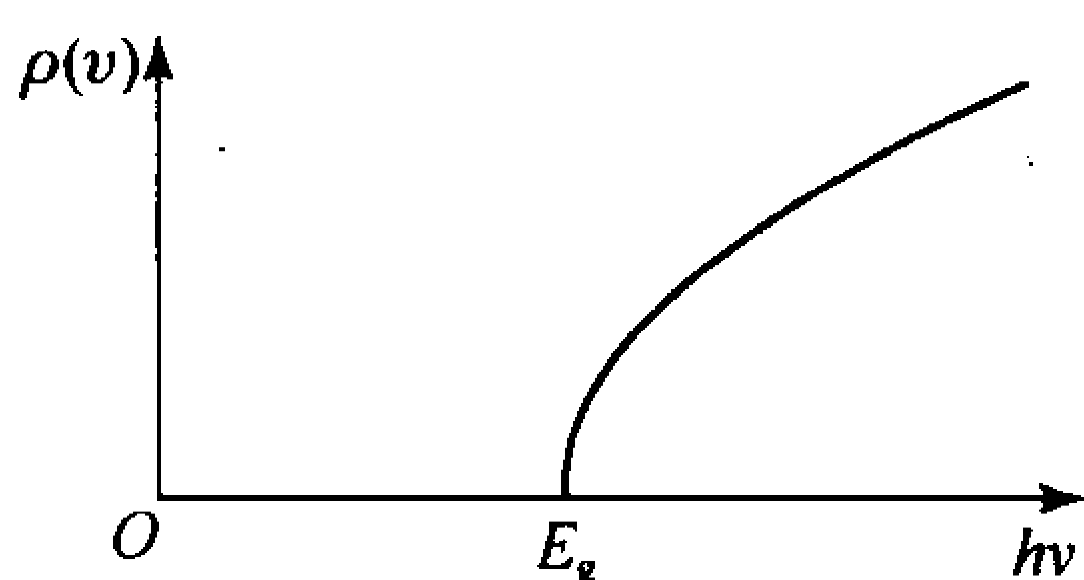


图 2.44 具有能量为 $h\nu$ 的光子与态密度的关系

(5) 间接带隙半导体中不太可能发生光子发射. 电子和空穴的辐射复合在一个间接带隙半导体中不太可能发生. 这是因为从导带底附

近到价带顶附近(导带底和价带顶分别是电子和空穴最可能存在的地方)的跃迁需要一个动量的交换,这个动量不能够由发射出去的光子提供,然而动量可以通过多声子的参与而守恒,如图 2.45 所示. 声子可以携带相对较大的动量,然而在一般典型情况下只能携带很小的能量(约为 0.01～0.1 eV,参见图 2.41),因此声子的跃迁基本上是横向的. 其净结果就是动量守恒了,但是违反了 k 选择定则. 因为声子辅助的发射包含了三体(电子、光子、声子)参与的相互作用,所以该作用发生的几率十分小. 这样,作为间接带隙半导体的 Si 的辐射复合比作为直接带隙半导体的 GaAs 要少. 因此,Si 不是高效的发光器,而 GaAs 却是.

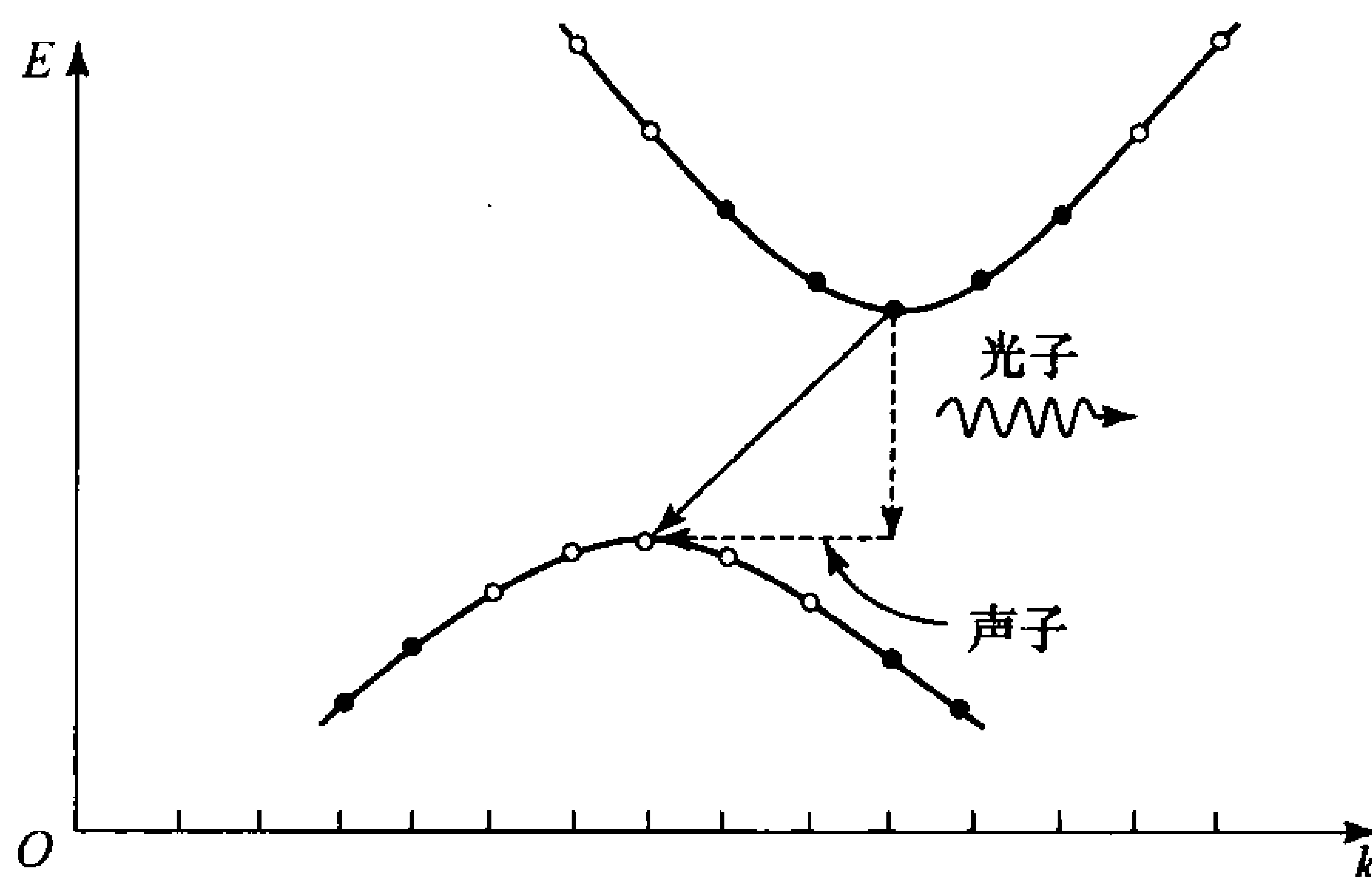


图 2.45　间接带隙半导体中的光子发射

(6) 间接带隙半导体中的光子吸收是可能的. 尽管在间接带隙半导体中光子吸收也需要能量和动量的守恒,但这容易通过两步过程来实现(图 2.46). 电子首先通过一个垂直的跃迁被激发到导带中的一个高能量的能级. 然后,电子通过一个叫做热化的过程迅速退激发到导带底. 在这个过程中,电子的动量被传递给了声子. 吸收光子产生的空穴有着相似的行为. 因为这个过程是连续发生的,

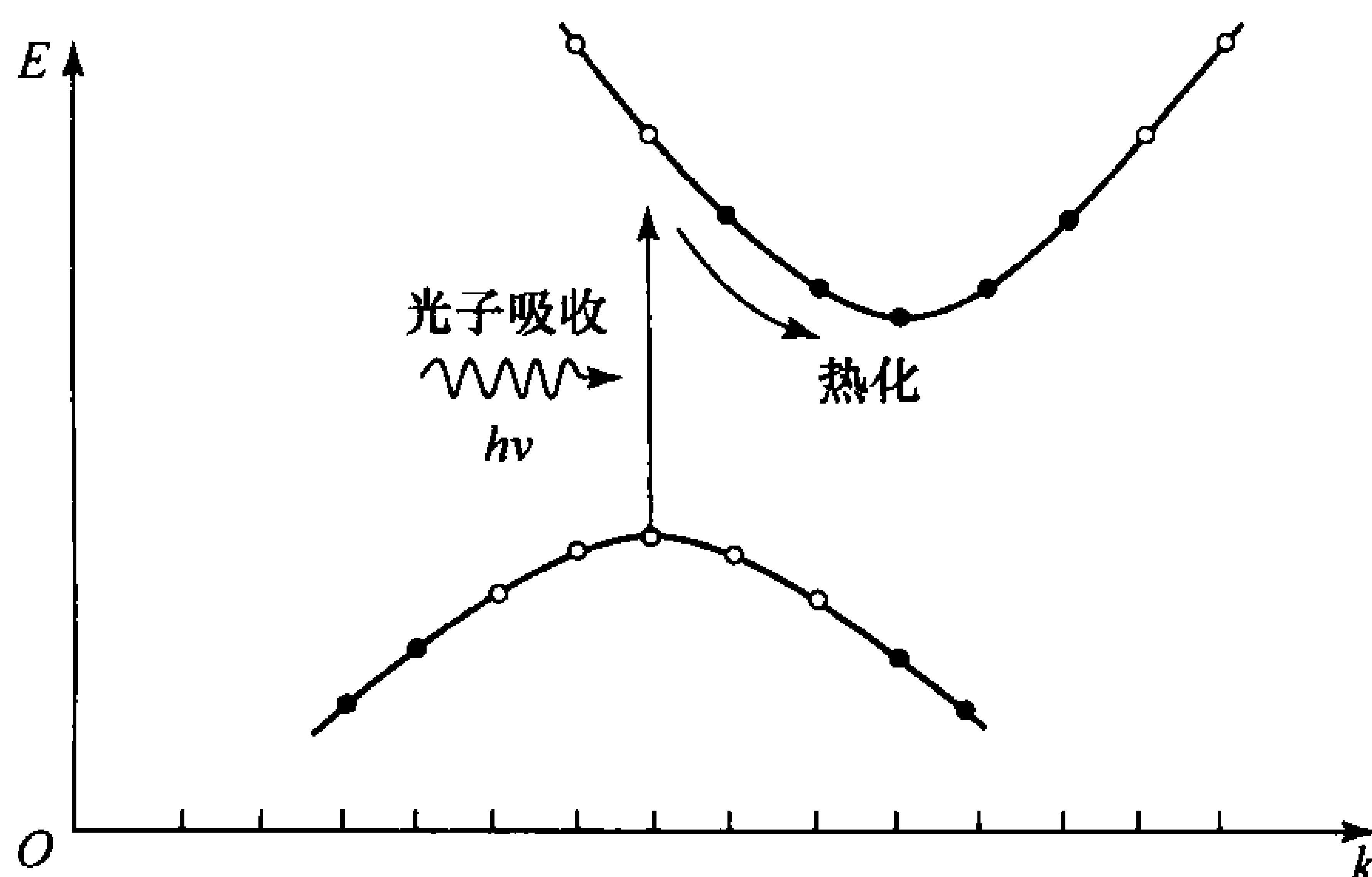


图 2.46　间接带隙半导体中的光子吸收

所以它不需要三种粒子同时参与相互作用，因此是可能发生的. 所以 Si 可以是一个高效的光探测器，就像 GaAs 一样.

2.4.2 吸收和发射速率

本小节将确定在直接带隙跃迁过程中一个能量为 $h\nu$ 的光子被半导体材料所发射或者吸收的概率密度. 用以描述能量动量守恒的表达式(2.4.6)，(2.4.7)和(2.4.4)决定了与光子发生相互作用的电子和空穴的能量 E_1，E_2 及动量 $\hbar k$.

有三个因素决定了这个概率密度：占据概率、跃迁概率和态密度. 我们按顺序讨论之.

1. 占据概率

通过在不连续的能级 E_1 和 E_2 间的跃迁从而实现光子发射和吸收的占据条件如下：

(1) 发射条件：一个导带的能量状态被（一个电子）占据，而一个价带的能量状态是空着的（即被空穴占据）；

(2) 吸收条件：一个导带的能量状态是空着的并且一个价带的能量状态被占据.

这些占据条件在各种 E_1 和 E_2 取值下被满足的概率是由适当的费米函数 $f_c(E)$ 和 $f_v(E)$ 决定的. 这些函数对应于半导体的导带和价带中的准平衡过程. 因此，对一个能量为 $h\nu$ 的光子，发射条件被满足的概率 $f_e(\nu)$ 是上能级被占据同时下能级空着的两个独立事件的概率，即

$$f_e(\nu) = f_c(E_2)[1 - f_v(E_1)], \tag{2.4.10}$$

E_1 和 E_2 通过式(2.4.6)和(2.4.7)和 ν 相关联. 类似地，吸收条件被满足的概率 $f_a(\nu)$ 为

$$f_a(v) = [1 - f_c(E_2)]f_v(E_1). \tag{2.4.11}$$

2. 跃迁概率

满足发射（吸收）占据条件并不能保证发射（吸收）能真正发生. 这些过程也被光子和原子系统相互作用的概率法则所支配. 当它们和半导体相关时，这些法则一般涉及发射光子到 $\nu\sim\nu+d\nu$ 之间的一个狭窄频带（或者从该狭窄频带中吸收光子），如下所述：

一次在两个不连续的能级 E_1 和 E_2 之间的辐射跃迁可以由跃迁横截面 $\sigma(\nu)=(\lambda^2/8\pi\tau_{sp})g(\nu)$ 来描述，其中 ν 是频率，τ_{sp} 是自发的寿命，而 $g(\nu)$ 是谱线形状函数（具有单位面积，线宽为 $\Delta\nu$，以跃迁频率 $\nu_0=(E_2-E_1)/h$ 为中心的函数）. 在半导体中，电子空穴辐射复合寿命 τ_r 扮演了 τ_{sp} 的角色，从而有

$$\sigma(\nu) = \frac{\lambda^2}{8\pi\tau_r}g(\nu). \tag{2.4.12}$$

如果发射光子的占据条件被满足，单位时间内自发辐射一个光子到 $\nu \sim \nu + \mathrm{d}\nu$ 这一狭窄频带中任一可能的辐射模式的概率密度为

$$P_{sp}(\nu)\mathrm{d}\nu = \frac{1}{\tau_r} g(\nu)\mathrm{d}v. \tag{2.4.13}$$

如果发射光子的占据条件满足并且在频率 ν 处有一个平均的光子流密度 Φ_ν（单位时间内单位面积上单位频率的光子），则在单位时间内自发辐射一个光子到 $\nu \sim \nu + \mathrm{d}\nu$ 之间的狭窄频带的概率密度为

$$W_i(\nu)\mathrm{d}\nu = \Phi_\nu \sigma(\nu)\mathrm{d}\nu = \Phi_\nu \frac{\lambda^2}{8\pi\tau_r} g(\nu)\mathrm{d}\nu. \tag{2.4.14}$$

如果吸收光子的占据条件满足并且在频率 ν 处有一个平均的光子流密度 Φ_ν，则从 $\nu \sim \nu + \mathrm{d}\nu$ 之间的狭窄频带吸收一个光子的概率密度也由式(2.4.14)表示.

由于每一次跃迁都会有一个不同的中心频率 ν_0，而且我们要考虑许多这类跃迁，为清楚起见，我们把 $g(\nu)$ 记为 $g_{\nu_0}(\nu)$，以区分跃迁的中心波长. 在半导体中，与两个能级相关，且各向均匀展宽的谱线形状函数一般都来源于电子-光子碰撞展宽. 因此它一般表现出洛仑兹线型，其展宽 $\Delta\nu \approx 1/(\pi T_2)$，其中电子空穴的碰撞时间 T_2 是在皮秒量级. 例如如果 $T_2 = 1$ ps，则 $\Delta\nu = 318$ GHz，相应的能量展宽为 $h\Delta\nu \approx 1.3$ meV. 能级的辐射寿命展宽相对于碰撞展宽来说是可以忽略的.

3. 发射和吸收速率

在一对由 $E_2 - E_1 = h\nu_0$ 分离的能级之间，自发发射速率、受激发射速率以及对于能量为 $h\nu$ 的光子的吸收速率（半导体中单位时间内单位体积上单位频率的光子数，单位 $\mathrm{s}^{-1} \cdot \mathrm{cm}^{-3} \cdot \mathrm{Hz}^{-1}$）在频率为 ν 的情况下可以由以下得到. 以合适的跃迁概率密度 $P_{sp}(\nu)$ 或者 $W_i(\nu)$ [在式(2.4.13)和(2.4.14)中给出]乘以合适的占据概率 $f_e(\nu_0)$ 或者 $f_a(\nu_0)$[由式(2.4.10)或者(2.4.11)给出]，再乘以与光子相互作用的态密度 $\rho(\nu_0)$[由式(2.4.9)给出]. 因此，对于所有允许的频率 ν_0，其总的跃迁概率可以通过对 ν_0 积分而得到.

例如对于频率为 ν 的自发发射速率，可以由下式给出：

$$r_{sp}(\nu) = \int [(1/\tau_r) g_{\nu_0}(\nu)] f_e(\nu_0)\rho(\nu_0)\mathrm{d}\nu_0.$$

通常情况下，当碰撞展宽 $\Delta\nu$ 实际上比函数 $f_e(\nu_0)\rho(\nu_0)$ 的宽度小时，$g\nu_0(\nu)$ 可以近似地用 $\delta(\nu - \nu_0)$ 表示，于是自发发射速率被简化为

$$r_{sp}(\nu) = (1/\tau_r)\rho(\nu) f_e(\nu). \tag{2.4.15}$$

受激发射速率和吸收速率可以由同样的方法得到，因此有以下公式

$$r_{st}(\nu) = \Phi_\nu \frac{\lambda^2}{8\pi\tau_r}\rho(\nu) f_e(\nu), \tag{2.4.16}$$

$$r_{ab}(\nu) = \Phi_\nu \frac{\lambda^2}{8\pi\tau_r}\rho(\nu) f_a(\nu). \tag{2.4.17}$$

这些公式加上式(2.4.9),(2.4.10)和(2.4.11),使得来自带间跃迁的自发发射、受激发射和吸收速率(单位时间内单位体积上单位频率的光子数,单位 $s^{-1}\cdot cm^{-3}\cdot Hz^{-1}$)可以通过平均光子流密度 Φ_ν(单位时间内单位面积上单位频率,单位 $s^{-1}\cdot cm^{-2}\cdot Hz^{-1}$)来计算. 而乘积 $\rho(\nu)f_e(\nu)$ 和 $\rho(\nu)f_a(\nu)$ 分别类似于上、下两个能级的谱线形状函数和原子数密度的乘积,即 $g(\nu)N_2$ 和 $g(\nu)N_1$.

占据概率 $f_e(\nu)$ 和 $f_a(\nu)$ 的确定需要知道准费米能级 E_{f_c} 和 E_{f_v}. 人们正是通过控制这两个参数(比如利用给 p-n 节加电压的手段)来控制修改发射和吸收的速率,从而制成具有不同功能的半导体光子器件. 式(2.4.15)描述了 LED 的工作原理;式(2.4.16)适用于半导体光放大器和注入激光器;式(2.4.17)适用于基于光子吸收工作的半导体光探测器.

4. 热平衡下的自发辐射谱密度

处于热平衡下的半导体只有一个费米函数,因此式(2.4.10)变成 $f_e(\nu)=f(E_2)[1-f(E_1)]$. 如果费米能级处于禁带之中,远离带边至少几个 k_BT,则可以用指数函数来近似描述费米函数,即

$$f(E_2)\approx\exp[-(E_2-E_f)/k_BT],$$
$$1-f(E_1)\approx\exp[-(E_f-E_1)/k_BT],$$

而 $f_e(\nu)\approx\exp[-(E_2-E_1)/k_BT]$,即

$$f_e(\nu)\approx\exp\left(-\frac{h\nu}{k_BT}\right).\tag{2.4.18}$$

把式(2.4.15)中的 $\rho(\nu)$ 用式(2.4.10)代入,$f_e(\nu)$ 用式(2.4.18)代入,则可以得到

$$r_{sp}(\nu)\approx D_0(h\nu-E_g)^{1/2}\exp\left(-\frac{h\nu-E_g}{k_BT}\right),\quad h\nu\geqslant E_g,\tag{2.4.19}$$

其中

$$D_0=\frac{(2m_r)^{3/2}}{\pi\hbar^2\tau_r}\exp\left(-\frac{E_g}{k_BT}\right)\tag{2.4.20}$$

是一个与温度有指数关系的参量.

图 2.47 中描绘了自发发射速率与 $h\nu$ 的关系曲线,有两个因素影响这一关系:① 从态密度角度来考虑的按 $h\nu-E_g$ 幂次增加的关系;② 从费米函数角度来影响的按 $h\nu-E_g$ 负指数衰减的关系.

自发发射速率可以通过增加 $f_e(\nu)$ 来增加. 如式(2.4.10)所描述,通过使材料偏离热平衡,可以增大 $f_c(E_2)$ 且减小 $f_v(E_1)$ 较小. 这样可以确保材料中具有足够的电子和空穴,而这些也正是使得 LED 正常工作所需要的.

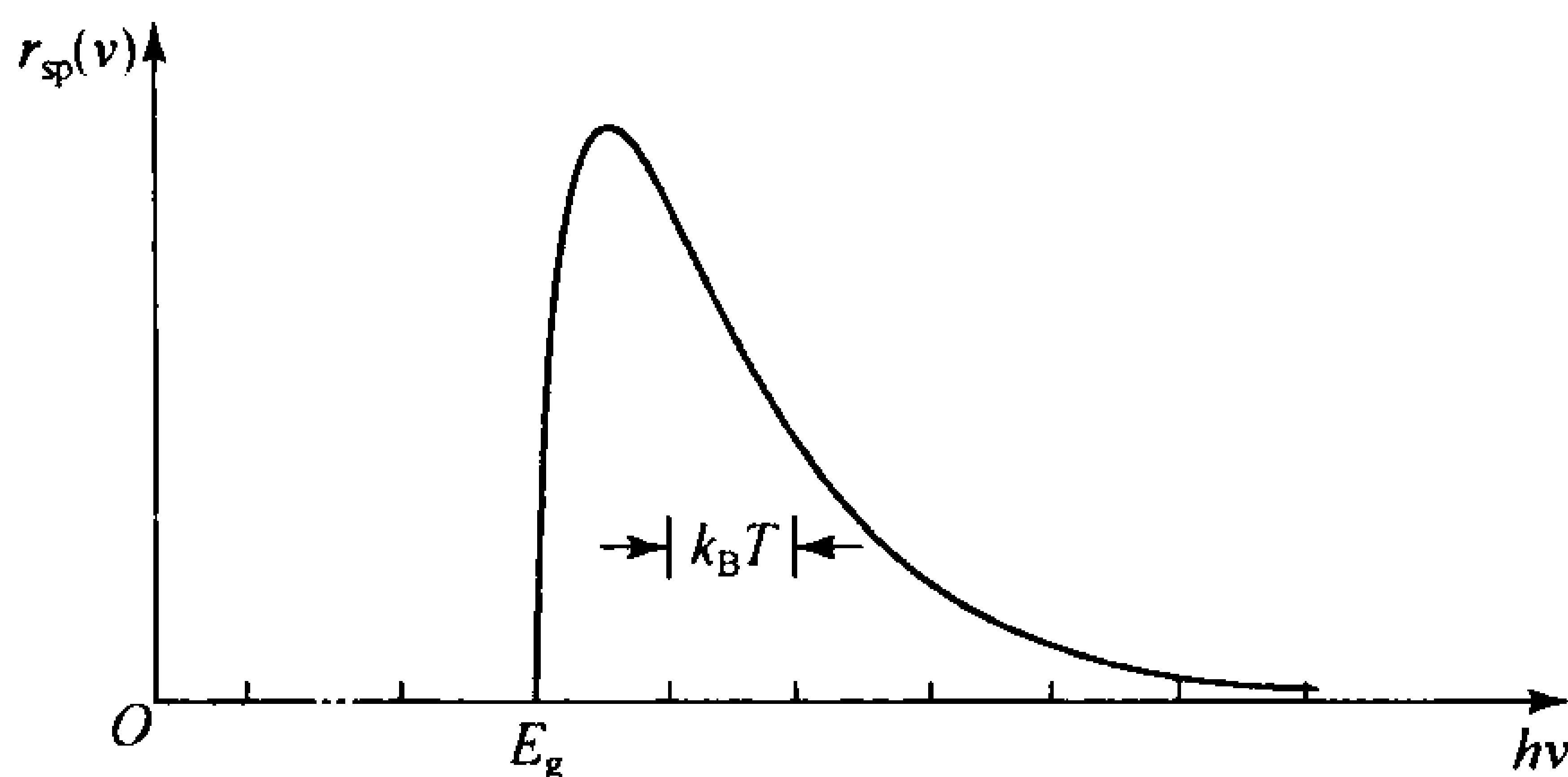


图 2.47 处于热平衡的半导体中，直接带间自发发射率（单位时间内单位体积上单位频率的光子数，单位 $s^{-1}\cdot cm^{-3}\cdot Hz^{-1}$）的谱密度作为 $h\nu$ 的函数曲线. 谱线宽度为 $2k_BT/h$ 且有一个较低的截止频率 $\nu=E_g/h$

5. 准平衡下的增益系数

受激发射速率[参见式(2.4.16)]和吸收速率[参见式(2.4.17)]的净增益系数 $r_0(\nu)$ 可以由如下方法确定. 选取一个单位面积的圆柱和长度增量 dz，假设平均光子流的谱密度是随 z 轴变化的函数. 如果 $\Phi_\nu(z)$ 和 $\Phi_\nu(z)+d\Phi_\nu(z)$ 分别是进入和离开圆柱的平均光子流谱密度，则 $d\Phi_\nu(z)$ 必为从圆柱发射出来的平均光子流密度谱密度. 由于单位时间内单位面积上单位频率增加的光子数是单位时间单位频率单位体积的光子数增益 $r_{st}(\nu)-r_{ab}(\nu)$ 乘以圆柱体的厚度 dz，即

$$d\Phi_\nu(z)=[r_{st}(\nu)-r_{ab}(\nu)]dz,$$

利用式(2.4.16)和(2.4.17)，我们得到

$$\frac{d\Phi_\nu(z)}{dz}=\frac{\lambda^2}{8\pi\tau_r}\rho(\nu)[f_e(\nu)-f_a(\nu)]\Phi_\nu(z)=r_0(\nu)\Phi_\nu(z). \quad (2.4.21)$$

因此，净增益系数为

$$r_0(\nu)=\frac{\lambda^2}{8\pi\tau_r}\rho(\nu)f_g(\nu), \quad (2.4.22)$$

其中费米反转因子由下式表示：

$$f_g(\nu)\equiv f_e(\nu)-f_a(\nu)=f_c(E_2)-f_v(E_1), \quad (2.4.23)$$

这可以从式(2.4.10)和(2.4.11)，以及通过式(2.4.6)和(2.4.7)描述的 E_1 及 E_2 同 ν 的依赖关系看出. 我们利用式(2.4.9)，则增益系数可以由下述形式计算：

$$r_0(\nu)=D_1(h\nu-E_g)^{1/2}f_g(\nu), \quad h\nu>E_g,$$

其中

$$D_1=\frac{\sqrt{2}m_r^{3/2}\lambda^2}{h^2\tau_r}.$$

费米反转因子 $f_g(\nu)$ 的符号和频谱形状是由准费米能级 E_{f_c} 和 E_{f_v} 决定，它

们又反过来依赖于半导体中载流子的激发状态. 这个因子只有在 $E_{f_c}-E_{f_v}>h\nu$ 时才是正的(对应一个净的粒子数反转). 当通过外界能源使半导体被激发到一个非常高的态时,这个条件才可能满足从而得到净增益. 这些也正是半导体光放大器和注入激光器得以运行的物理基础.

6. 热平衡下的吸收系数

处于热平衡下的半导体只有一个费米能级 $E_f=E_{f_c}=E_{f_v}$,因此

$$f_c(E)=f_v(E)=f(E)=\frac{1}{\exp[(E-E_f)/k_BT]+1}. \quad (2.4.24)$$

由于反转因子 $f_g(\nu)=f_c(E_2)-f_v(E_1)=f(E_2)-f(E_1)<0$,因此净增益系数 $r_0(\nu)$总是负的[因为 $E_2>E_1$ 并且 $f(E)$随 E 增加而单调减少],无论 E_f 是在什么位置这都是成立的. 所以,一个处于热平衡的半导体无论是本征半导体还是掺杂半导体,总是使光衰减. 衰减系数(或者吸收系数)$\alpha(\nu)=-r_0(\nu)$为

$$\alpha(\nu)=D_1(h\nu-E_g)^{1/2}[f(E_1)-f(E_2)], \quad (2.4.25)$$

其中 E_1 和 E_2 由式(2.4.7)和(2.4.6)分别给出,而 D_1 由式(2.4.23b)给出.

如果 E_f 处于带隙中但距离带边至少有几个 k_BT 的距离,则 $f(E_1)\approx1$, $f(E_2)\approx0$,因此 $f(E_1)-f(E_2)\approx1$. 在这种情况下,直接带间跃迁对吸收系数的贡献为

$$\alpha(\nu)\approx\frac{\sqrt{2}c^2m_r^{3/2}}{\tau_r}\frac{1}{(h\nu)^2}(h\nu-E_g)^{1/2}. \quad (2.4.26)$$

当温度上升时,$f(E_1)-f(E_2)$下降至小于 1 并且吸收系数也减少. 图 2.48 描绘 GaAs 遵照式(2.4.26)的变化规律,其中用到了下列参数: $n=3.6$, $m_c=0.07m_0$, $m_v=0.5m_0$, $m_0=9.1\times10^{-31}$ kg,使得 $\tau_r=0.4$ ns 的掺杂浓度, $E_g=1.42$ eV,以及使得 $f(E_1)-f(E_2)\approx1$ 的温度.

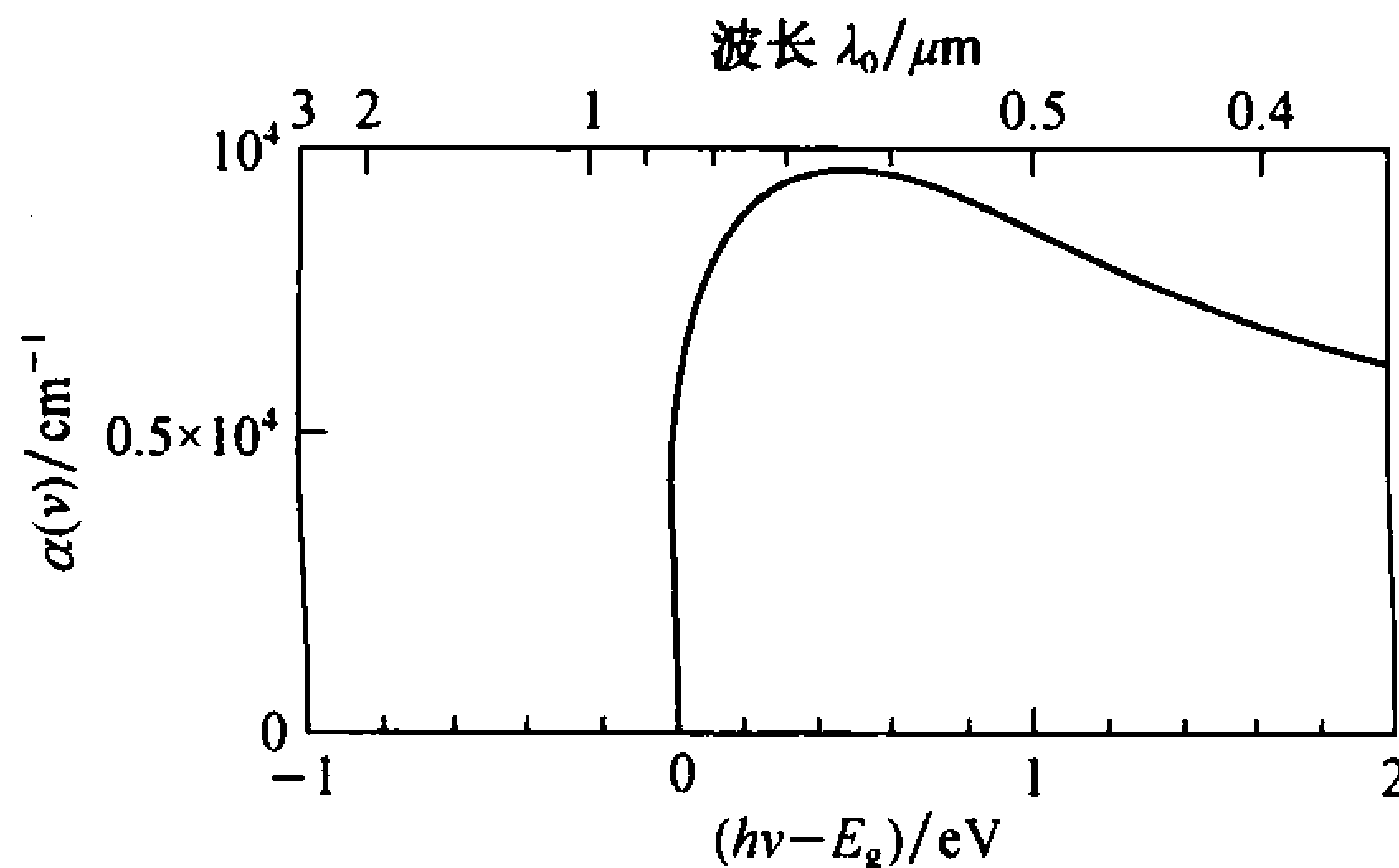


图 2.48 通过直接带间跃迁计算得出的 GaAs 吸收系数作为光子能量和波长的函数曲线;与图 2.41 的实验结果相比较,后者包括了所有的吸收机制

2.4.3　折射率

在光学和光子学中，折射率是研究最多的参数之一. 可以说，任何光学现象及对它们的控制，都或多或少与折射率的变化有关. 因此，控制半导体的折射率对光电子器件的设计是非常重要的. 半导体材料是色散材料，因此折射率依赖于波长. 实际上，折射率和吸收系数 $\alpha(\nu)$ 的依赖关系就像极化率的实部及虚部必须满足克拉默斯-克勒尼希(Kramers-Kronig, K-K)关系一样. 更为重要的是，折射率同样也依赖于温度和掺杂情况. 图 2.49 显示了 GaAs 材料中，折射率随波长和掺杂浓度的变化而变化的规律.

半导体材料折射率随掺杂浓度的变化使得人们能够利用微电子的制作手段及方法方便地制作出以半导体材料为平台的光学(光电子器件)及系统. 硅基光电子学的核心任务之一就是研究如何利用硅或与硅工艺兼容的其他材料，开发以光子、电子为载体的微纳量级信息功能器件，并将它们在同一硅基底上通过 CMOS 工艺进行大规模集成，形成一个完整的具有综合功能的新型芯片单元. 显然，理解和应用 CMOS 工艺，并利用它来有选择性调控硅材料的折射率是硅基光电子学的重要基础之一.

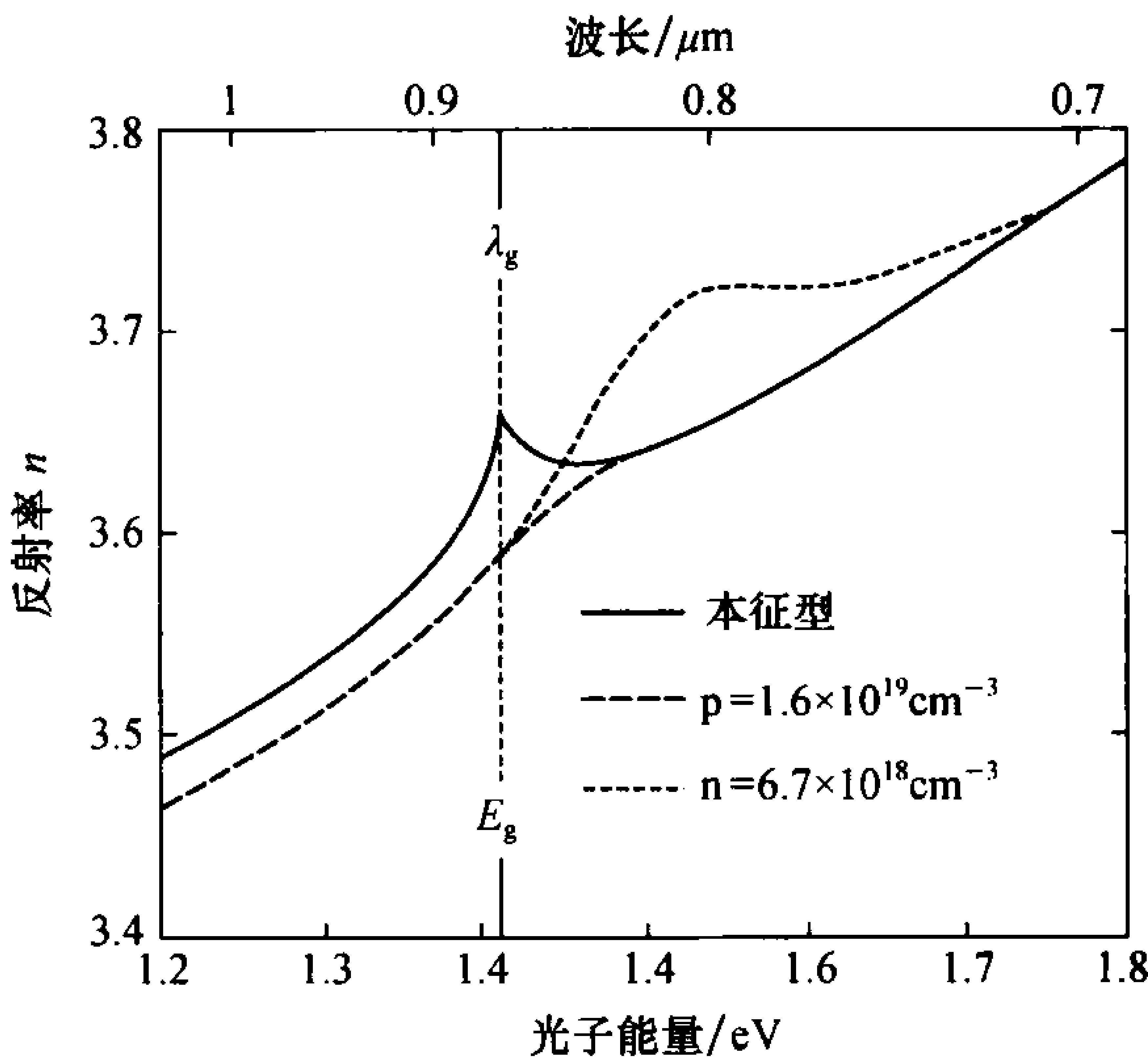


图 2.49　高纯度、p 型掺杂和 n 型掺杂的 GaAs 在 300K 下的折射率作为光子能量(波长)的函数图；在高纯样品的曲线中，带隙波长处的峰值是与自由激子相关联的[32]

作为比较，表 2.4 给出了在特定情况下和接近带隙波长处，部分一元和二元半导体在 300 K 下的折射率.

表 2.4 当光子波长与带隙波长相当时,部分半导体材料在室温下的折射率

材料	折射率
半导体元素	
Ge	4.0
Si	3.5
Ⅲ-Ⅴ族二元半导体	
AlP	3.0
AlAs	3.2
AlSb	3.8
GaP	3.3
GaAs	3.6
GaSb	4.0
InP	3.5
InAs	3.8
InSb	4.2

参考文献

[1] Saleh B E A, Teich M C. Fundamentals of Photonics. New Jersey: John Wiley & Sons Inc. ,1991.

[2] Yablonovitch E. Inhibited spontaneous emission in solid-state physics and electronics. Phys. Rev. Lett. , 1987, 58(20): 2059.

[3] John S. Strong localization of photons in certain disordered dielectric superlattices. Phys. Rev. Lett. , 1987, 58(23): 2486.

[4] Rowson S, Chelnokov A, Lourtioz J M. Two-dimensional photonic crystals in macroporous silicon: From mid-infrared (10 μm) to telecommunication wavelengths (1.3～1.5 μm). J. Lightwave Technology, 1999, 17(11): 1989.

[5] Netti M C, Charlton M D B, Parker G J, et al. Visible photonic band gap engineering in silicon nitride waveguides. Appl. Phys. Lett. , 2000, 76(8): 991.

[6] Birner A, Wehrspohn R B, Gosele U M, et al. Silicon-based photonic crystals. Adv. Mater. , 2001, 13(6): 377.

[7] Chow E, Lin S Y, Johnson S G, et al. Three-dimensional control of light in a two-dimensional photonic crystal slab. Nature, 2000, 407(6807): 983.

[8] Qiu M. Band gap effects in asymmetric photonic crystal slabs. Phys. Rev. B, 2002, 66(3): 033103.

[9] Johnson S G, Fan S H, Villeneuve P R, et al. Guided modes in photonic crystal slabs. Phys. Rev. B, 1999, 60(8): 5751.

[10] Rinne S A, Garcia S F, Braun P V. Embedded cavities and waveguides in three-

dimensional silicon photonic crystals. Nature Photonics, 2008, 2(1): 52.

[11] Kurt H, Citrin D S. Annular photonic crystals. Opt. Exp., 2005, 13(25): 10316.

[12] Wen F, David S, Checoury X, et al. Two-dimensional photonic crystals with large complete photonic band gaps in both TE and TM polarizations. Opt. Exp., 2008, 16 (16): 12278.

[13] Shi P, Huang K, Kang X L, et al. Creation of large band gap with anisotropic annular photonic crystal slab structure. Opt. Exp., 2010, 18(5): 5221.

[14] Marsal L F, Trifonov T, Rodriguez A, et al. Larger absolute photonic band gap in two-dimensional air-silicon structures. Physica E-Low-Dimensional Systems and Nanostructures, 2003, 16(3—4): 580.

[15] Poborchii V, Tada T, Kanayama T, et al. Silver-coated silicon pillar photonic crystals: Enhancement of a photonic band gap. Appl. Phys. Lett., 2003, 82(4): 508.

[16] Toader O, John S. Photonic band gap enhancement in frequency-dependent dielectrics. Phys. Rev. E, 2004, 70(4): 6605.

[17] Fu H K, Chen Y F, Chern R L, et al. Connected hexagonal photonic crystals with largest full band gap. Opt. Exp., 2005, 13(20): 7854.

[18] Abrarov S M, Abrarov R M. Broadening of band-gap in photonic crystals with optically saturated media. Opt. Commu., 2008, 281(11): 3131.

[19] Estevez J O, Arriaga J, Blas A M, et al. Enlargement of omnidirectional photonic bandgap in porous silicon dielectric mirrors with a Gaussian profile refractive index. Appl. Phys. Lett., 2009, 94(6): 1914.

[20] Meade R D, Rappe A M, Brommer K D, et al. Nature of the photonic band-gap-some insights from a field Analysis. J. Opt. Soc. Am. B: Opt. Phys., 1993, 10(2): 328.

[21] Akahane Y, Asano T, Song B S, et al. Fine-tuned high-Q photonic-crystal nanocavity. Opt. Exp., 2005, 13(4): 1202.

[22] Kurt H, Hao R, Chen Y, et al. Design of annular photonic crystal slabs. Opt. Lett., 2008, 33(14): 1614.

[23] Saynatjoki A, Mulot M, Vynck K, et al. Properties, applications and fabrication of photonic crystals with ring-shaped holes in silicon-on-insulator. Photonics and Nanostructures: Fundamentals and Applications, 2008, 6(1): 42.

[24] Cicek A, Ulug B. Polarization-independent waveguiding with annular photonic crystals. Opt. Exp., 2009, 17(20): 18381.

[25] Jagerska J, Le T N, Zabelin V, et al. Experimental observation of slow mode dispersion in photonic crystal coupled-cavity waveguides. Opt. Lett., 2009, 34(3): 359.

[26] Chang K P, Yang S L, Shen L F, et al. Photonic band gaps of polygonal and circular dielectric rods in square lattices. J. Appl. Phys., 2008, 104(11): 3109.

[27] Hou J, Gao D S, Wu H M, et al. Polarization insensitive self-collimation waveguide in square lattice annular photonic crystals. Opt. Commu., 2009, 282(15): 3172.

[28] Mulot M, Saynatjoki A, Arpiainen S, et al. Slow light propagation in photonic crystal waveguides with ring-shaped holes. J. Opt. A: Pure Appl. Opt. , 2007, 9(9): S415.

[29] Feng J B, Chen Y, Blair J, et al. Fabrication of annular photonic crystals by atomic layer deposition and sacrificial etching. J. Vac. Sci. and Tech. B, 2009, 27(2): 568.

[30] Qiu M. Effective index method for heterostructure-slab-waveguide-based two-dimensional photonic crystals. Appl. Phys. Lett. , 2002, 81(7): 1163.

[31] Srinivasan K, Painter O. Momentum space design of high-Q photonic crystal optical cavities. Opt. Exp. , 2002, 10(15): 670.

[32] Casey H C and Panish M B. Heterostructure Lasers Part A: Fundamental Principles. New York: Academic Press, 1978.

第三章　硅基光波导

§3.1　电磁理论基础

19 世纪 60 年代，伟大的英国物理学家麦克斯韦在前人工作的基础上，尤其是在法拉第、高斯等人对电磁现象进行的深入研究的基础上，加上他的天才假设，提出了完整描述宏观电磁现象的一套基本方程，这就是麦克斯韦方程组[1~3]. 根据这组基本方程，麦克斯韦预言了电磁波的存在，并指出光波就是波长极短的电磁波，从而圆满解决了光波的本质这一悬而未决的问题，翻开了光的经典电磁理论这一物理学史上的新篇章. 迄今为止，除了光的发射与吸收必须用量子理论才能圆满解释以外，麦克斯韦的经典电磁理论仍然是分析光的传输问题的理论基础. 本章将从麦克斯韦方程组出发推出光波传播的一系列基本规律[4~7].

3.1.1　电磁场基本方程

1. 麦克斯韦方程组

宏观电磁现象可以用电场矢量 $\boldsymbol{E}$、电位移矢量 $\boldsymbol{D}$、磁场强度矢量 $\boldsymbol{H}$、磁感应强度矢量 $\boldsymbol{B}$ 等四个矢量描述，它们都是空间位置和时间的函数. 这四个场矢量的关系由麦克斯韦方程描述，即

$$\begin{cases} \nabla\times \boldsymbol{H} = \boldsymbol{J} + \dfrac{\partial \boldsymbol{D}}{\partial t}, \\ \nabla\times \boldsymbol{E} = -\dfrac{\partial \boldsymbol{B}}{\partial t}, \\ \nabla\cdot \boldsymbol{B} = 0, \\ \nabla\cdot \boldsymbol{D} = \rho, \end{cases} \tag{3.1.1}$$

式中 $\boldsymbol{J}$ 是媒质的传导电流密度，ρ 是自由电荷密度. 式(3.1.1)中的四个方程式不是独立的，如果认为电流连续性方程

$$\nabla\cdot \boldsymbol{J} + \frac{\partial \rho}{\partial t} = 0 \tag{3.1.2}$$

是独立方程，则式(3.1.1)中后两个方程可以由前两个方程式推出. 要从式(3.1.1)中完全确定电磁场量，尚需给出 $\boldsymbol{D}$,$\boldsymbol{B}$ 与 $\boldsymbol{E}$,$\boldsymbol{H}$ 之间的关系，这一组关系式称为本

构关系或物质特性方程，即

$$\begin{cases} \boldsymbol{J} = \sigma\boldsymbol{E}, \\ \boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P}, \\ \boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M}), \end{cases} \tag{3.1.3}$$

式中 $\boldsymbol{P}$ 称为媒质的极化强度矢量，$\boldsymbol{M}$ 称为磁化强度矢量. σ 是媒质的电导率，对良好的介质可以认为近似为零，ε_0，μ_0 分别为真空的介电常数和磁导率. 对于非磁性介质，$\boldsymbol{M}=\boldsymbol{0}$，从而

$$\boldsymbol{B} = \mu_0\boldsymbol{H}.$$

电极化强度 $\boldsymbol{P}$ 可以写为

$$\boldsymbol{P} = \varepsilon_0\boldsymbol{\chi}^{(1)} \cdot \boldsymbol{E} + \varepsilon_0\boldsymbol{\chi}^{(2)} : \boldsymbol{EE} + \varepsilon_0\boldsymbol{\chi}^{(3)} ⋮ \boldsymbol{EEE} + \cdots, \tag{3.1.4}$$

其中 $\boldsymbol{\chi}^{(i)}$ 是 $i+1$ 阶张量，

$$\boldsymbol{\chi}^{(i)} = \begin{bmatrix} \chi & 0 & 0 \\ 0 & \chi & 0 \\ 0 & 0 & \chi \end{bmatrix}.$$

如果除去 $\boldsymbol{\chi}^{(1)}$ 以外，所有的 $\boldsymbol{\chi}^{(i)}$ 的元都为零，则媒质是线性媒质，否则媒质是非线性的. 对于各向同性的线性媒质，它可以用一个标量 χ 表示，从而得到

$$\begin{cases} \boldsymbol{P} = \varepsilon_0\chi\boldsymbol{E}, \\ \boldsymbol{D} = \varepsilon_0(1+\chi)\boldsymbol{E} = \varepsilon_0\varepsilon_r\boldsymbol{E}, \quad \varepsilon_r = (1+\chi). \end{cases} \tag{3.1.5}$$

一般说来，所有的场量都是空间坐标和时间的任意函数，一个时域函数可以用傅里叶变换表示：

$$\begin{cases} \boldsymbol{\Psi}(x,y,z,t) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} \boldsymbol{\Psi}(x,y,z,\omega)\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}\omega, \\ \boldsymbol{\Psi}(x,y,z,\omega) = \displaystyle\int_{-\infty}^{\infty} \boldsymbol{\Psi}(x,y,z,t)\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}t, \end{cases} \tag{3.1.6}$$

式中 $\boldsymbol{\Psi}(x,y,z,t)$ 可代表所有场分量的时域表达式，而 $\boldsymbol{\Psi}(x,y,z,\omega)$ 则为其频域表达式.

在良好的介质中，电流密度和电荷密度都为零，于是频域中的麦克斯韦方程组可以写成

$$\begin{cases} \nabla\times\boldsymbol{H} = \mathrm{i}\omega\varepsilon_0\varepsilon_r\boldsymbol{E}, \\ \nabla\times\boldsymbol{E} = -\mathrm{i}\omega\mu_0\boldsymbol{H}, \\ \nabla\cdot\boldsymbol{H} = 0, \\ \nabla\cdot\varepsilon_0\varepsilon_r\boldsymbol{E} = 0. \end{cases} \tag{3.1.7}$$

2. 电磁场边界条件

麦克斯韦方程组(3.1.1)描述的是电磁参数 ε，μ 为位置坐标的连续函数的

媒质中电磁场量的基本规律[8,9]. 如果媒质的电磁参数发生突变,则以微分方程形式出现的麦克斯韦方程组将不再适用,此时需将方程组(3.1.1)改写成它的积分形式,即

$$\oint_l \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{l}=\iint_S J\cdot \mathrm{d}\boldsymbol{S}+\iint_S \frac{\partial \boldsymbol{D}}{\partial t}\cdot \mathrm{d}\boldsymbol{S}, \tag{3.1.8a}$$

$$\oint_l \boldsymbol{E}\cdot \mathrm{d}\boldsymbol{l}=-\iint_S \frac{\partial \boldsymbol{D}}{\partial t}\cdot \mathrm{d}\boldsymbol{S}, \tag{3.1.8b}$$

$$\oiint_S \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{S}=0, \tag{3.1.8c}$$

$$\oiint_S \boldsymbol{D}\cdot \mathrm{d}\boldsymbol{S}=\iiint_V \rho \mathrm{d}V. \tag{3.1.8d}$$

上式中的 $\boldsymbol{E}$,$\boldsymbol{H}$ 在封闭曲线 l 上线积分,其积分路径即为式(3.1.8a),(3.1.8b)右边面积分区域 S 的边界,而式(3.1.8c)和(3.1.8d)左边 $\boldsymbol{B}$ 和 $\boldsymbol{D}$ 在封闭曲面上的面积分的积分区域,即为右边体积分区域 V 的外表面. 将式(3.1.8c)和(3.1.8d)应用于如图 3.1(a)所示的扁平区域,可得

$$\boldsymbol{n}\cdot(\boldsymbol{B}_1-\boldsymbol{B}_2)=0, \tag{3.1.9a}$$

$$\boldsymbol{n}\cdot(\boldsymbol{D}_1-\boldsymbol{D}_2)=\rho_S. \tag{3.1.9b}$$

式(3.1.9a)说明,在两种媒质分界面的两侧磁感应强度 $\boldsymbol{B}$ 的法向分量连续,面电位移矢量 $\boldsymbol{D}$ 的法向分量的突变决定于界面上的面电荷密度 ρ_S.

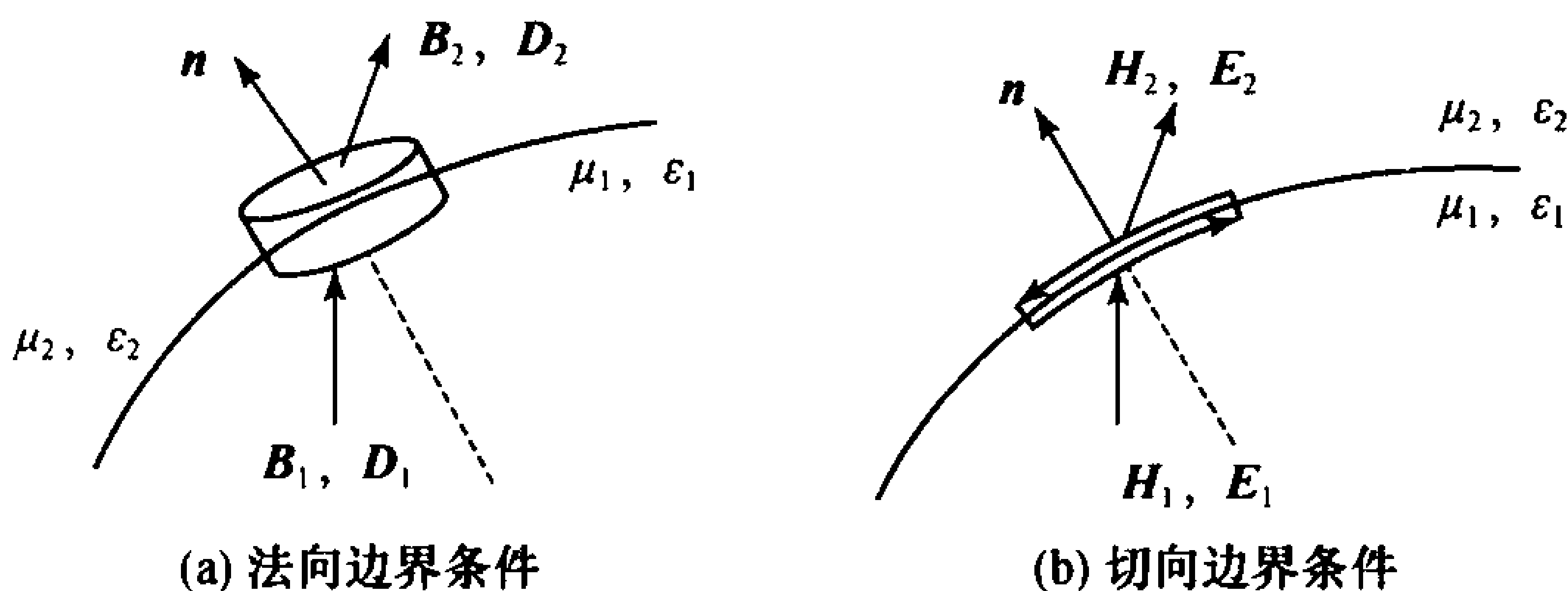


图 3.1　确定电磁场边界条件的几何区域

如果将式(3.1.8a)和(3.1.8b)应用于图 3.1(b)所示的窄条形路径,则可得到

$$\boldsymbol{n}\times(\boldsymbol{H}_1-\boldsymbol{H}_2)=\boldsymbol{J}_S, \tag{3.1.10a}$$

$$\boldsymbol{n}\times(\boldsymbol{E}_1-\boldsymbol{E}_2)=\boldsymbol{0}. \tag{3.1.10b}$$

式(3.1.10a)说明,磁场强度 $\boldsymbol{H}$ 的切向分量在边界面上的突变取决于界面上的面电流密度 $\boldsymbol{J}_S$,而电场强度 $\boldsymbol{E}$ 的切向分量则总是连续的. 可以认为面电流仅存在于理想导体表面,因而总可以认为,实际的两种媒质分界面两侧磁场强度 $\boldsymbol{H}$

的切向分量也是连续的.

对于非导电的介质,其表面电荷密度 $\rho_S=0$,表面电流密度 $\boldsymbol{J}_S=\boldsymbol{0}$,因而可以将(3.1.9)式和(3.1.10)式合并写成

$$\boldsymbol{n}\cdot(\boldsymbol{B}_1-\boldsymbol{B}_2)=0, \tag{3.1.11a}$$

$$\boldsymbol{n}\cdot(\boldsymbol{D}_1-\boldsymbol{D}_2)=0, \tag{3.1.11b}$$

$$\boldsymbol{n}\times(\boldsymbol{H}_1-\boldsymbol{H}_2)=\boldsymbol{0}, \tag{3.1.11c}$$

$$\boldsymbol{n}\times(\boldsymbol{E}_1-\boldsymbol{E}_2)=\boldsymbol{0}, \tag{3.1.11d}$$

式中 $\boldsymbol{n}$ 为界面上由媒质 1 指向媒质 2 的法线方向的单位矢量.

3. 波动方程和亥姆霍兹方程

在良好的介质中 $\boldsymbol{J}=\boldsymbol{0}$,$\rho=0$,如果媒质为均匀的各向同性的线性媒质,则 ε_r 为常数,在上述假设条件下,将方程组(3.1.1)中的前两个方程取旋度,并注意到

$$\nabla\cdot\boldsymbol{D}=\varepsilon_0\varepsilon_r\nabla\cdot\boldsymbol{E}=0,$$

$$\nabla\cdot\boldsymbol{B}=\mu_0\nabla\cdot\boldsymbol{H}=0,$$

则可得到

$$\nabla^2\boldsymbol{E}-\frac{n^2}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2}=0, \tag{3.1.12a}$$

$$\nabla^2\boldsymbol{H}-\frac{n^2}{c^2}\frac{\partial^2\boldsymbol{H}}{\partial t^2}=0, \tag{3.1.12b}$$

式中,$c=1/\sqrt{\mu_0\varepsilon_0}$是真空中的光速,$n=\sqrt{\varepsilon_r}$是媒质的折射率,式(3.1.12)即为线性、均匀、各向同性媒质中的电磁波动方程,它的解即为以速度 $v=c/n$ 传播的电磁波.

在频率域中,所有的场量都是以角频率 ω 震荡的正弦量,因而其波动方程则为

$$\nabla^2\boldsymbol{E}+k_0^2n^2\boldsymbol{E}=0, \tag{3.1.13a}$$

$$\nabla^2\boldsymbol{H}+k_0^2n^2\boldsymbol{H}=0, \tag{3.1.13b}$$

式中

$$k_0^2=\omega^2\mu_0\varepsilon_0. \tag{3.1.14}$$

式(3.1.13)称为亥姆霍兹方程,对于非均匀的各向同性线性媒质,则因为 $\nabla\cdot\boldsymbol{D}=\nabla\cdot\varepsilon_0\varepsilon_r\boldsymbol{E}=\varepsilon_0\nabla\varepsilon_r\boldsymbol{E}+\varepsilon_0\varepsilon_r\nabla\cdot\boldsymbol{E}=0$,可得

$$\nabla\cdot\boldsymbol{E}=-\frac{\nabla\varepsilon_r}{\varepsilon_r}\cdot\boldsymbol{E},$$

从而可得到

$$\nabla^2\boldsymbol{E}+k_0^2n^2\boldsymbol{E}+\nabla\left(\boldsymbol{E}\cdot\frac{\nabla\varepsilon_r}{\varepsilon_r}\right)=\boldsymbol{0}, \tag{3.1.15a}$$

$$\nabla^2\boldsymbol{H}+k_0^2n^2\boldsymbol{H}+\frac{\nabla\varepsilon_r}{\varepsilon_r}\times(\nabla\times\boldsymbol{H})=\boldsymbol{0}. \tag{3.1.15b}$$

式(3.1.15)中 $n^2=\varepsilon_r$ 是位置的函数,如果媒质的折射率或相对介电常数随位置的变化较为缓慢,即满足条件 $\left|\frac{\nabla\varepsilon_r}{\varepsilon_r}\right|\ll 1$,则称这种媒质为缓变媒质,于是式(3.1.15)可以简写成

$$\nabla^2\boldsymbol{E}+k_0^2n^2\boldsymbol{E}=0, \tag{3.1.16a}$$

$$\nabla^2\boldsymbol{H}+k_0^2n^2\boldsymbol{H}=0. \tag{3.1.16b}$$

式(3.1.16)虽然与式(3.1.13)形式上相同,但二者有着重要区别,即式(3.1.16)中的折射率 n 是空间位置的函数,因而其求解也就是要困难得多.

在分析光波导中光波的传输特点时,我们既会遇到均匀媒质,又会遇到非均匀媒质,但光波导中媒质的非均匀性总满足所谓缓变条件,因而式(3.1.13)和(3.1.16)是我们分析光波导中光波传播的基础[10~13].

4. 柱形波导中的场方程

柱型波导在光波系统中占有最重要的地位,光波在柱型波导中传播时,其电磁场量可以写成纵向分量和横向分量之和,即

$$\begin{cases}\boldsymbol{E}=\boldsymbol{E}_t+\boldsymbol{e}_zE_z,\\ \boldsymbol{H}=\boldsymbol{H}_t+\boldsymbol{e}_zH_z,\end{cases} \tag{3.1.17}$$

式中 $\boldsymbol{e}_z$ 为柱型系统中的纵向(也就是 z 轴正方向)的单位矢量.分析柱型系统的坐标系如图 3.2 所示.在图 3.2 所示的坐标系中,沿 z 轴方向传播的光波,其所有的场分量都必有 $\mathrm{e}^{-i\beta z}$ 这样的传播因子,β 为波的相位常数,因而亥姆霍兹方程(3.1.13)可以写成

$$\nabla_t^2\left[\frac{\boldsymbol{E}_t}{\boldsymbol{H}_t}\right]+(k_0^2n^2-\beta^2)\left[\frac{\boldsymbol{E}_t}{\boldsymbol{H}_t}\right]=\boldsymbol{0}, \tag{3.1.18a}$$

$$\nabla_t^2\left(\frac{E_z}{H_z}\right)+(k_0^2n^2-\beta^2)\left(\frac{E_z}{H_z}\right)=0, \tag{3.1.18b}$$

式中∇_t^2是拉普拉斯(Laplace)算子.对 z 向均匀的柱型传输系统,我们可以从矢量亥姆霍兹方程中分离出纵向分量 E_z,H_z 所满足的标量亥姆霍兹方程(3.1.18b).求解式(3.1.18b)得到纵向场量 E_z,H_z 以后,已无需求解式(3.1.18a).横向场量 $\boldsymbol{E}_t$ 和 $\boldsymbol{H}_t$ 可以从麦克斯韦方程(3.1.7)直接得到.

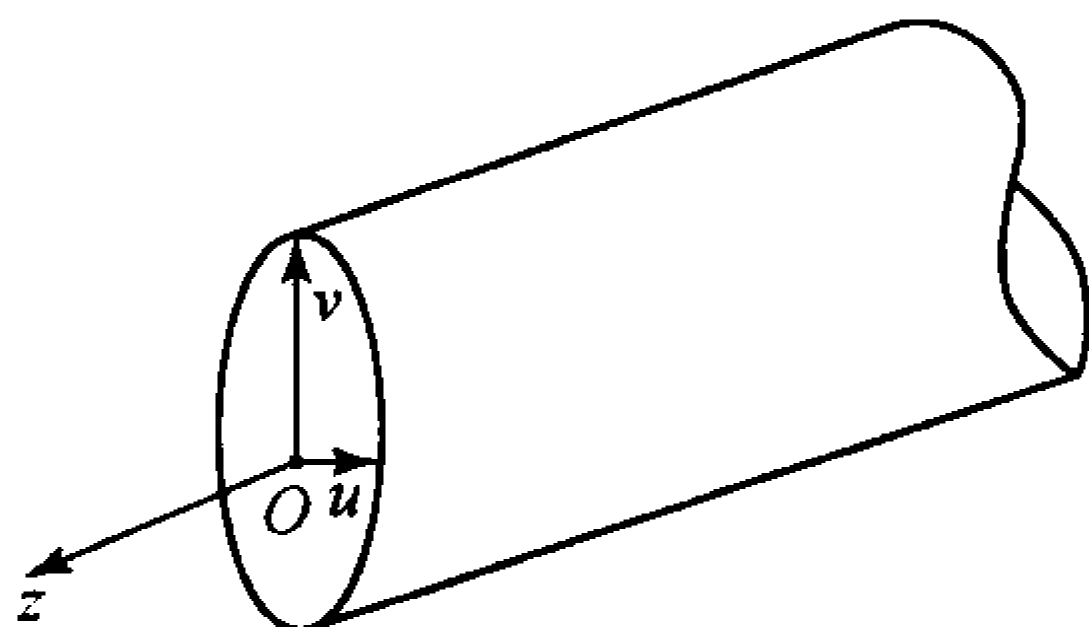


图 3.2　分析柱型系统所用的坐标系

在式(3.1.7)中把梯度算子写成$\nabla=\nabla_t+\boldsymbol{e}_z\dfrac{\partial}{\partial z}$,并将式(3.1.7)中的方程展开,即可得到

$$\begin{cases} k_0^2\boldsymbol{E}_t = -\mathrm{i}\beta\nabla_t E_z + \mathrm{i}\omega\mu_0\boldsymbol{e}_z\times\nabla_t H_z, & (3.1.19\mathrm{a})\\ k_\mathrm{e}^2\boldsymbol{H}_t = -\mathrm{i}\beta\nabla_t H_z + \mathrm{i}\omega\varepsilon_0\varepsilon_\mathrm{r}\boldsymbol{e}_z\times\nabla_t E_z, & (3.1.19\mathrm{b})\end{cases}$$

式中

$$k_\mathrm{e}^2 = k_0^2 n^2 - \beta^2. \tag{3.1.20}$$

3.1.2 各向同性媒质中的平面电磁波

1. 无界均匀媒质中的均匀平面电磁波

均匀平面电磁波的场量用下式表示:

$$\boldsymbol{E} = \boldsymbol{E}_0\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{3.1.21a}$$

$$\boldsymbol{H} = \boldsymbol{H}_0\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{3.1.21b}$$

式中$\boldsymbol{E}_0$和$\boldsymbol{H}_0$是波的振幅矢量,对均匀平面波它们都是常矢量,$\boldsymbol{k}$称为波矢量,它的方向即为波的传播方向,大小$k=k_0n$,即为波的相位常数. 令

$$\boldsymbol{k}\cdot\boldsymbol{r} = k_x x + k_y y + k_z z = C, \tag{3.1.22}$$

式中C是任意常数. 上式在空间描出一组平面,称为波的等相位面.

将式(3.1.21)代入式(3.1.7),可得

$$\boldsymbol{E}_0 = -\eta\boldsymbol{e}_n\times\boldsymbol{H}_0, \tag{3.1.23a}$$

$$\boldsymbol{H}_0 = -\frac{1}{\eta}\boldsymbol{e}_n\times\boldsymbol{E}_0, \tag{3.1.23b}$$

$$\boldsymbol{k}\cdot\boldsymbol{E}_0 = 0, \tag{3.1.23c}$$

$$\boldsymbol{k}\cdot\boldsymbol{H}_0 = 0, \tag{3.1.23d}$$

式中$\eta=\sqrt{\mu_0/\varepsilon}$是媒质的波阻抗. 上式说明,均匀无界媒质中的均匀平面电磁波是 TEM 波.

平面电磁波的相速度和群速度分别定义为

$$v_\mathrm{p} = \frac{\omega}{\beta},\quad v_\mathrm{g} = \frac{\mathrm{d}\omega}{\mathrm{d}\beta}. \tag{3.1.24}$$

对于均匀平面波,$\beta=k_0n=\omega\sqrt{\mu_0\varepsilon_0}\,n$. 如果媒质的折射率$n$是与频率无关的常数,则有

$$v_\mathrm{p} = v_\mathrm{g} = \frac{1}{n\sqrt{\mu_0\varepsilon_0}} = \frac{c}{n}. \tag{3.1.25}$$

也就是说,对均匀无色散的媒质,其中传播的平面电磁波的相速度和群速度相等,而且与频率无关.

2. 平面电磁波的偏振状态

平面电磁波的偏振状态是指电场强度矢量或磁场强度矢量的空间取向随时

间的变化情况. 如果对一个确定的观察点,场矢量始终在一个确定的方向上振动,矢量的尖端轨迹是一个线段,则称为线偏振波;如果场矢量的尖端的轨迹是一个圆,则称为圆偏振波;如果场矢量的尖端的轨迹是一个椭圆,则称为椭圆偏振波.

我们知道,任意的场矢量总可以写成沿两个特征方向的分矢量之和,即

$$\boldsymbol{E} = \boldsymbol{e}_1 E_1 \mathrm{e}^{\mathrm{i}\varphi_1} + \boldsymbol{e}_2 E_2 \mathrm{e}^{\mathrm{i}\varphi_2}, \tag{3.1.26}$$

式中 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 为波传播的横方向上两个相互正交的单位矢量,而且 $\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_n$, $\boldsymbol{e}_n$ 是波传播方向上的单位矢量,φ_1, φ_2 分别为两个分量的相位因子. 易于证明,当 $\delta = \varphi_1 - \varphi_2 = 0, \pi$ 时,为线偏振波;当 $\delta = \pi/2, 3\pi/2$,而且 $E_1 = E_2$ 时,为圆偏振波;当 $\delta = \pi/2$ 时,场矢量的旋转方向与波的传播方向呈右手螺旋关系,称为右旋圆偏振波;而 $\delta = 3\pi/2$ 时,则为左旋圆偏振波. 除上述两种情形以外,波呈椭圆偏振状态,$0 < \delta < \pi$ 时为右旋椭圆波,而 $\pi < \delta < 2\pi$ 时则为左旋椭圆波. 需要说明的是,在一般的光学教材中圆偏振和椭圆偏振波的旋向与我们这里的规定刚好相反,这里的定义与工程电磁理论中的规定一致.

式(3.1.26)说明,任何一种偏振状态的平面波都可以看成沿 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 方向偏振的两个有确定相位关系的线偏振波的叠加. 类似地,我们也可以将任意一个平面波看成是两个旋向相反的圆偏振波的叠加,即

$$\boldsymbol{E} = E_{\mathrm{L}}(\boldsymbol{e}_1 + \mathrm{i}\boldsymbol{e}_2) \pm E_{\mathrm{R}}(\boldsymbol{e}_1 - \mathrm{i}\boldsymbol{e}_2), \tag{3.1.27}$$

式中 $E_{\mathrm{L}}, E_{\mathrm{R}}$ 分别为左旋和右旋圆偏振波的振幅. 如果 $E_{\mathrm{L}} = E_{\mathrm{R}}$,则式(3.1.27)代表一个线偏振波;如果 $E_{\mathrm{L}} \neq E_{\mathrm{R}}$,则代表椭圆偏振波;如果 $E_{\mathrm{L}}, E_{\mathrm{R}}$ 中任意一个为零,则为圆偏振波.

引进斯托克斯参数来描述波的偏振状态是方便的,4 个斯托克斯参数的定义是

$$\begin{cases} S_0 = E_1^2 + E_2^2, \\ S_1 = E_1^2 - E_2^2, \\ S_2 = 2E_1 E_2 \cos(\varphi_2 - \varphi_1), \\ S_3 = 2E_1 E_2 \sin(\varphi_2 - \varphi_1). \end{cases} \tag{3.1.28}$$

这 4 个参数不是独立的,很显然

$$S_0^2 = S_1^2 + S_2^2 + S_3^2. \tag{3.1.29}$$

如果以 S_1, S_2, S_3 与直角坐标系中的 x, y, z 坐标对应,则当 S_0 为常数时,决定了一个球面,这个球面称为庞加莱(Poincaré)球. 庞加莱球在描述波的偏振态方面极为有用,因为任何偏振状态都与邦加球面上的点有一一对应的关系,如果测得斯托克斯参数的变化规律,则可确定光波的偏振态变化规律,这为测定单模光纤传输系统的偏振模色散提供了理论基础.

3. 平面波的反射和折射

平面电磁波在不同媒质的平面分界面上将发生反射和折射，如图 3.3 所示. 根据分界面两侧电磁场应满足的边界条件，可得两种介质中的入射波、反射波、折射波之间的如下运动学特性：

① 入射光线、反射光线和折射光线共面，或波矢量 $\boldsymbol{k}_i,\boldsymbol{k}_r,\boldsymbol{k}_t$ 共面；

② 反射角等于入射角，即 $\theta_i=\theta_r$；

③ 菲涅耳定律：$n_1\sin\theta_i=n_2\sin\theta_t$，这里的 n_1,n_2 分别为两种媒质的折射率.

根据入射波的偏振状态，可以得到如下动力学性质：

① 对于垂直偏振波，即电场矢量与入射面垂直的线偏振波，其反射系数和折射系数分别为

$$\begin{cases} R_\perp = \dfrac{E_r}{E_i} = \dfrac{n_1\cos\theta_i - n_2\cos\theta_t}{n_1\cos\theta_i + n_2\cos\theta_t}, & (3.1.30a) \\ T_\perp = \dfrac{E_t}{E_i} = \dfrac{2n_1\cos\theta_i}{n_1\cos\theta_i + n_2\cos\theta_t}. & (3.1.30b) \end{cases}$$

② 对于平行偏振波，即电场矢量与入射面平行的线偏振波，其反射系数和折射系数分别为

$$\begin{cases} R_{/\!/} = \dfrac{E_r}{E_i} = \dfrac{n_2\cos\theta_i - n_1\cos\theta_t}{n_2\cos\theta_i + n_1\cos\theta_t}, & (3.1.31a) \\ T_{/\!/} = \dfrac{E_t}{E_i} = \dfrac{2n_1\cos\theta_i}{n_2\cos\theta_i + n_1\cos\theta_t}. & (3.1.31b) \end{cases}$$

式(3.1.30)和(3.1.31)中的 4 个关系式统称为菲涅耳公式.

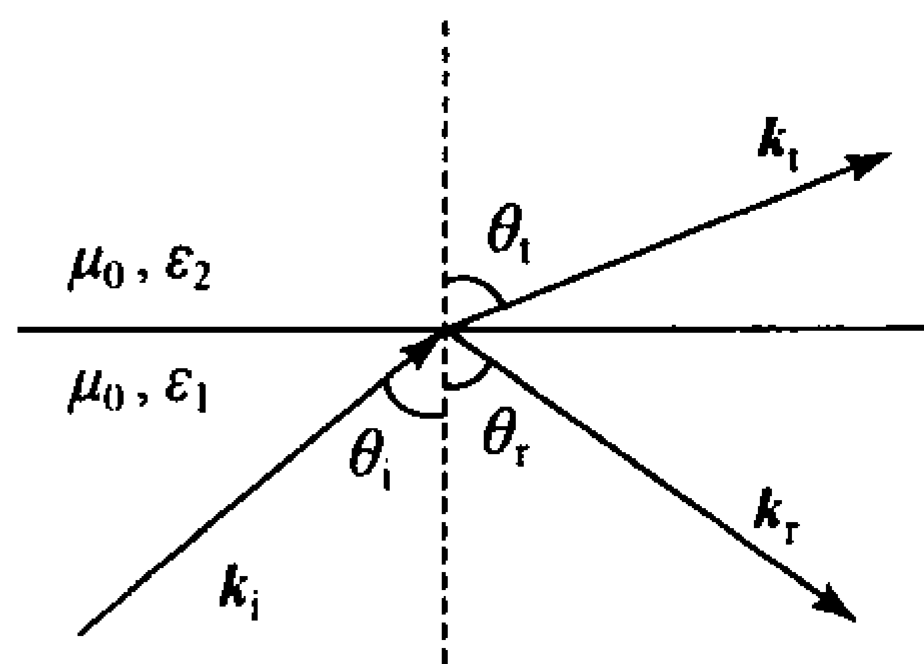


图 3.3 平面电磁波的反射和折射

由菲涅耳定律可知，当 $n_1>n_2$ 时，折射角大于入射角. 如果入射角 θ_i 等于临界角 θ_c，使得 $\sin\theta_t=1$，即 $\theta_t=\pi/2$，折射光线将与界面平行. 如果 $\theta_i>\theta_c$ 则折射光线消失从而产生全反射. 平面波从媒质 1 到媒质 2 的界面上产生全反射的条件是

$$\theta_i > \theta_c = \arcsin\frac{n_2}{n_1}. \qquad (3.1.32)$$

显然，仅当 $n_1>n_2$ 时，才有可能发生全反射. 如果入射条件满足

$$\theta_B = \arctan\frac{n_2}{n_1}, \tag{3.1.33}$$

则平行偏振波的反射系数为零，因而以 θ_B 入射的平行偏振波产生全折射. θ_B 称为布儒斯特角，又称为起偏振角，因为不管入射波的偏振状态如何，反射波都是线偏振波.

4. 非理想媒质中的平面电磁波

实际的媒质，除真空以外，都不是理想媒质. 实际媒质的电磁参数 ε，μ 都不是实数，而且电导率 $\sigma\neq0$，因而实际媒质中电磁波传播时都有损耗，而且媒质的折射率 n 都是频率的函数，从而导致色散. 实际媒质中平面波传播的两个最重要的特性就是损耗特性和色散特性.

实际媒质中电磁波的传播问题变得十分复杂. 对光波系统，我们关心的是损耗极低、色散极弱的良好介质中的电磁波，在这种介质中，沿 z 轴方向的平面波可以表示为

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{-\alpha z} e^{-ik_0 nz}, \tag{3.1.34a}$$

$$\boldsymbol{H} = \frac{\boldsymbol{E}_0}{\eta} e^{-\alpha z} e^{-ik_0 nz}, \tag{3.1.34b}$$

式中 α 为衰减常数，而 $k_0 n(\omega)=\beta$ 则为其相位常数.

对于非理想介质中的平面波，其相速度和群速度则为

$$v_p = \frac{\omega}{\beta} = \frac{c}{n(\omega)}, \tag{3.1.35a}$$

$$v_g = \frac{d\omega}{d\beta} = \frac{c}{n(\omega)+\omega\dfrac{dn(\omega)}{d\omega}}. \tag{3.1.35b}$$

为了说明群速度的物理意义，假设有一被调制的电磁信号在色散媒质中传播，将此电磁信号看成是沿 z 轴方向传播的但在 x-y 平面内均匀的脉冲信号，假设损耗可以忽略，则此信号可以表示为

$$A(z,t) = \int_{-\infty}^{\infty} A(\beta)\exp[i(\omega t-\beta(\omega)z)]d\beta. \tag{3.1.36}$$

式(3.2.36)说明一个空间长度为有限的信号可以在频域中表示为单色平面波的叠加，式中 $A(\beta)$ 是与 ω，$\beta(\omega)$ 相对应的振幅函数，显然

$$A(\beta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} A(z,t=0)\exp[i\beta(\omega)z]dz. \tag{3.1.37}$$

对色散媒质，$\beta(\omega)$ 与 ω 之间具有复杂的函数关系，将 $\omega(\beta)$ 在载波相位常数 β_0 附近展开为泰勒级数：

$$\omega(\beta) = \omega_0 + \frac{d\omega}{d\beta}(\beta-\beta_0) + \frac{1}{2}\frac{d^2\omega}{d\beta^2}(\beta-\beta_0)^2 + \cdots. \tag{3.1.38}$$

将其代入式(3.1.36)，忽略与$\frac{\mathrm{d}^2\omega}{\mathrm{d}\beta^2}$成比例的项以及更高阶的项，可得到

$$A(z,t)\approx \exp\left[\mathrm{i}\left(\omega_0 t-\frac{\mathrm{d}\omega}{\mathrm{d}\beta}\beta_0 z\right)\right]\int_{-\infty}^{\infty}A(\beta)\exp\left[\mathrm{i}\left(\frac{\mathrm{d}\omega}{\mathrm{d}\beta}t-z\right)\beta\right]\mathrm{d}\beta. \tag{3.1.39}$$

将上式中的积分与式(3.1.36)比较，可知此积分刚好就是$A\left(z-\frac{\mathrm{d}\omega}{\mathrm{d}\beta}t,0\right)$，因而可以得到

$$A(z,t)=A(z-v_g t,0)\exp[\mathrm{i}(\omega_0 t-v_g\beta_0 z)]. \tag{3.1.40}$$

从上式可以得到，已调制信号沿z轴传播时其信号包络以群速度$v_g=\frac{\mathrm{d}\omega}{\mathrm{d}\beta}$传播，在忽略与$\frac{\mathrm{d}^2\omega}{\mathrm{d}\beta^2}$成比例的项的情形下，信号包络形状不变，所以群速度也就是信号包络的传播速度.

3.1.3 各向异性媒质中的平面电磁波

1. 电各向异性媒质

媒质在电磁场作用下，根据其电极化和磁极化特性可以将媒质区分为各向同性、电各向异性、磁各向异性和双各向异性等几种类型. 日常所说的各向同性媒质只是一种近似，几乎所有的媒质都呈各向异性. 所谓电各向异性是指其介电常数为二阶张量，如大多数晶体；而磁各向异性则是指其磁导率为二阶张量，如外磁场作用下的铁氧体；所有的媒质在运动状态下，根据狭义相对论都呈双各向异性. 在光波系统中，经常会遇到光在晶体中的传播问题，因而我们主要关心的是电各向异性媒质中平面波的传播问题.

电各向异性媒质的介电常数，必须用一个二阶张量表示，即

$$\hat{\boldsymbol{\varepsilon}}=\begin{bmatrix}\varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13}\\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23}\\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33}\end{bmatrix}. \tag{3.1.41}$$

对于无损耗的电各向异性媒质，利用能量守恒定律可以证明式(3.1.41)中的张量为一个实对称张量，即$\varepsilon_{ij}=\varepsilon_{ji}$. 根据线性代数理论，任何一个实对称矩阵都存在一个正交的线性变换，可以将此实对称矩阵变换成一个对角阵. 这个线性变换实际上相当于一个坐标系的一个旋转. 在新的坐标系下，也就是所谓主轴坐标系中，介电常数张量可表示为

$$\hat{\boldsymbol{\varepsilon}}=\begin{bmatrix}\varepsilon_x & 0 & 0\\ 0 & \varepsilon_y & 0\\ 0 & 0 & \varepsilon_z\end{bmatrix}. \tag{3.1.42}$$

如果 $\varepsilon_x=\varepsilon_y\neq\varepsilon_z$，则称媒质为单轴媒质，$z$ 轴为其光轴；如果 $\varepsilon_x\neq\varepsilon_y\neq\varepsilon_z$，则称为双轴媒质；在 $\varepsilon_x=\varepsilon_y=\varepsilon_z$ 时，则称为各向同性媒质.

对电各向异性媒质，在主轴坐标系中

$$\boldsymbol{D}=\hat{\boldsymbol{\varepsilon}}\boldsymbol{E},$$

写成分量表示则为

$$D_x=\varepsilon_x E,\quad D_y=\varepsilon_y E,\quad D_z=\varepsilon_z E,$$

由此可以定义主轴坐标系中的折射率

$$n_i^2=\frac{\varepsilon_i}{\varepsilon_0},\quad i=x,y,z. \tag{3.1.43}$$

2. 电各向异性媒质中的平面波

各向异性媒质中的均匀平面电磁波，所有的场矢量均可表示为 $\boldsymbol{A}=\boldsymbol{A}_0\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$，$\boldsymbol{A}_0$ 为振幅矢量，将其代入麦克斯韦方程

$$\begin{cases}\nabla\times\boldsymbol{H}=\mathrm{i}\omega\boldsymbol{D},\\ \nabla\times\boldsymbol{E}=-\mathrm{i}\omega\mu_0\boldsymbol{H},\\ \nabla\cdot\boldsymbol{B}=0,\\ \nabla\cdot\boldsymbol{D}=0,\end{cases}$$

可得

$$\begin{cases}\boldsymbol{k}\times\boldsymbol{H}=-\omega\boldsymbol{D}, & (3.1.44\mathrm{a})\\ \boldsymbol{k}\times\boldsymbol{E}=\omega\mu_0\boldsymbol{H}, & (3.1.44\mathrm{b})\\ \boldsymbol{k}\cdot\boldsymbol{H}=0, & (3.1.44\mathrm{c})\\ \boldsymbol{k}\cdot\boldsymbol{D}=0. & (3.1.44\mathrm{d})\end{cases}$$

上式说明，各向异性媒质中的平面波是关于场矢量 $\boldsymbol{D}$ 和 $\boldsymbol{B}$ 的 TEM 波，但由于 $\boldsymbol{D}$ 与 $\boldsymbol{E}$ 不平行，所以不一定是关于 $\boldsymbol{E}$ 和 $\boldsymbol{H}$ 的 TEM 波.

为简单起见，我们讨论单轴媒质中的平面波. 将式(3.1.44b)左乘矢量 $\boldsymbol{k}$，并利用式(3.1.44a)，则有

$$\boldsymbol{k}\times\boldsymbol{k}\times\boldsymbol{E}=\omega\mu_0\boldsymbol{k}\times\boldsymbol{H}=-\omega^2\mu_0\boldsymbol{D},$$

或

$$\boldsymbol{k}(\boldsymbol{k}\cdot\boldsymbol{E})-k^2\boldsymbol{E}=-\omega^2\mu_0\hat{\boldsymbol{\varepsilon}}\cdot\boldsymbol{E}.$$

对单轴媒质，在主轴坐标系中

$$\hat{\boldsymbol{\varepsilon}}=\begin{bmatrix}\varepsilon & 0 & 0\\ 0 & \varepsilon & 0\\ 0 & 0 & \varepsilon_z\end{bmatrix},$$

因而

$$\boldsymbol{k}\cdot\boldsymbol{D}=\boldsymbol{k}\cdot\hat{\boldsymbol{\varepsilon}}\cdot\boldsymbol{E}=\varepsilon k_xE_x+\varepsilon k_yE_y+\varepsilon_zk_zE_z,$$

所以有

$$\boldsymbol{k}\cdot\boldsymbol{E}=\left(1-\frac{\varepsilon_z}{\varepsilon}\right)k_zE_z,$$

从而得到

$$\boldsymbol{k}\left(1-\frac{\varepsilon_z}{\varepsilon}\right)k_zE_z-k^2\boldsymbol{E}=-\omega^2\mu_0\hat{\boldsymbol{\varepsilon}}\cdot\boldsymbol{E},\tag{3.1.45}$$

将上式用矩阵表示则为

$$\begin{bmatrix}\omega^2\mu_0\varepsilon-k^2 & 0 & k_xk_z\left(1-\frac{\varepsilon_z}{\varepsilon}\right)\\ 0 & \omega^2\mu_0\varepsilon-k^2 & k_yk_z\left(1-\frac{\varepsilon_z}{\varepsilon}\right)\\ 0 & 0 & \omega^2\mu_0\varepsilon_z-k_x^2-k_y^2-k_z^2\end{bmatrix}\begin{bmatrix}E_x\\E_y\\E_z\end{bmatrix}=0.\tag{3.1.46}$$

显然，要使式(3.1.46)中的 E_x,E_y,E_z 有非零解的必要条件是与上述方程系数矩阵相应的行列式为零，从而得到

$$k^2=\omega^2\mu_0\varepsilon,\tag{3.1.47a}$$

$$k_x^2+k_y^2+\frac{\varepsilon_z}{\varepsilon}k_z^2=\omega^2\mu_0\varepsilon_z.\tag{3.1.47b}$$

式(3.1.47)即为单轴媒质中平面波传播的特征方程. 在 $\varepsilon=\varepsilon_z$ 时，上述两个方程即为同一个方程：$k^2=\omega^2\mu_0\varepsilon$，这就是我们熟知的各向同性媒质中平面波传播的特征方程. 显然在各向同性媒质中，平面波相位常数与 $\boldsymbol{k}$ 的方向无关，因而在任何方向上，平面波以相同的速度传播.

在单轴媒质中，由式(3.1.47a)所决定的相位常数 $k=\omega\sqrt{\mu_0\varepsilon_0}n=k_0n$ 与 $\boldsymbol{k}$ 的方向无关，也就是说，由式(3.1.47a)决定的特征相位常数及相应的平面波与各向同性媒质中的平面波具有相同的特点，因而称这种平面波为寻常波(o 波). 而由式(3.1.47b)所决定的特征常数则与波矢量 $\boldsymbol{k}$ 的方向有关，因而称这第二个因方向而异的特征相位常数所决定的平面波为非寻常波(e 波).

假设 $\boldsymbol{k}$ 与光轴 z 之间的夹角 θ，则可将式(3.1.47)写成

$$k^2\left(\sin^2\theta+\frac{\varepsilon_z}{\varepsilon}\cos^2\theta\right)=\omega^2\mu_0\varepsilon_z,\tag{3.1.48a}$$

或

$$k^2=\omega^2\mu_0\varepsilon_z\varepsilon/(\varepsilon\sin^2\theta+\varepsilon_z\cos^2\theta),\tag{3.1.48b}$$

从而得到寻常波和非寻常波的相速度分别为

$$\begin{cases}v_o=\dfrac{1}{\sqrt{\mu_0\varepsilon}}=\dfrac{c}{n_0}, & (3.1.49a)\\ v_e=\left(\dfrac{\varepsilon\sin^2\theta+\varepsilon_z\cos^2\theta}{\mu_0\varepsilon_z\varepsilon}\right)^{\frac{1}{2}}=\dfrac{c}{n_0n_e}(n_0^2\sin^2\theta+n_e^2\cos^2\theta)^{\frac{1}{2}}. & (3.1.49b)\end{cases}$$

由式(3.1.49b)可见，非寻常波的相速度 v_e 与 θ 有关，仅当 $\theta=0$ 时，即传播

方向与光轴一致时 v_e 才与寻常波的相速度 v_o 一致.

将式(3.1.47a)代入式(3.1.46),可得 $E_z=0$,从而得 $D_z=0$. 因为 $\varepsilon_x=\varepsilon_y=\varepsilon$,所以 $\boldsymbol{D}$ 与 $\boldsymbol{E}$ 同方向,显然 $\boldsymbol{B}$ 与 $\boldsymbol{H}$ 同方向,所以寻常波既是关于 $\boldsymbol{D},\boldsymbol{B}$ 的 TEM 波,也是关于 $\boldsymbol{E},\boldsymbol{H}$ 的 TEM 波. 由于 $\boldsymbol{E},\boldsymbol{D}$ 同时垂直于 $\boldsymbol{k}$ 和 z 轴,所以 $\boldsymbol{k}$ 与 z 轴不同方向时,$\boldsymbol{E},\boldsymbol{D}$ 垂直于 $\boldsymbol{k}$ 与 z 轴形成的平面,这种情形下寻常波只能是线偏振波. 如果 $\boldsymbol{k}$ 与 z 轴同方向,则 $\boldsymbol{E}$ 和 $\boldsymbol{D}$ 可以在 Oxy 平面内任意取向,因而寻常波可以是线偏振波、圆偏振波或椭圆偏振波.

对非寻常波,将式(3.1.47b)入式(3.1.46)可以决定场矢量的偏振状态. 不失一般性,假设 $\boldsymbol{k}$ 矢量在 Oyz 平面内,由于非寻常波的特点仅由 θ 决定,其他子午面内的情形与 Oyz 平面内是完全一样的. 在这种假设下

$$\boldsymbol{k} = \boldsymbol{e}_y k\sin\theta + \boldsymbol{e}_z k\cos\theta,$$

于是式(3.1.46)中的头两个方程为

$$\begin{cases} E_x = 0, \\ \varepsilon\sin\theta E_y + \varepsilon_z\cos\theta E_z = 0. \end{cases} \tag{3.1.50}$$

由此可知,电场矢量 $\boldsymbol{E}$ 也在 Oyz 平面内,且两个分量 E_z,E_y 同相位,所以非寻常波是线偏振波. 由于 $\boldsymbol{D}$ 与 $\boldsymbol{E}$ 不同方向,所以非寻常波已不是 $\boldsymbol{E},\boldsymbol{H}$ 意义下的 TEM 波. 寻常波和非寻常波的偏振状态分别如图 3.4 所示.

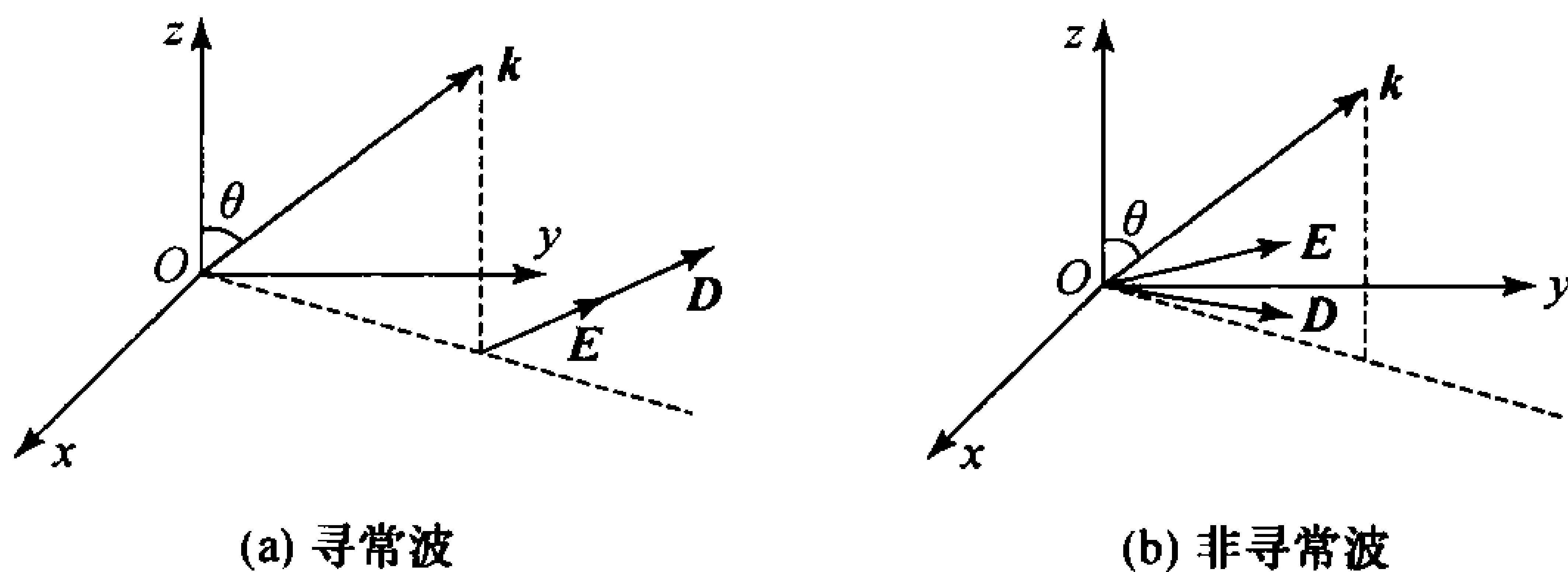


图 3.4　寻常波和非寻常波的偏振状态

由式(3.1.44)可知,在各向异性媒质中,$\boldsymbol{k}$ 与 $\boldsymbol{D}\times\boldsymbol{B}$ 同方向. $\boldsymbol{k}$ 矢量的方向代表平面波等相位面的传播方向,而 $\boldsymbol{D}\times\boldsymbol{B}$ 是与电磁场的动量相联系的量,所以 $\boldsymbol{k}$ 总与电、磁动量矢量同方向. 由于 $\boldsymbol{E}$ 与 $\boldsymbol{D}$ 不同方向,因而 $\boldsymbol{E}\times\boldsymbol{H}$ 与 $\boldsymbol{D}\times\boldsymbol{B}$ 不同方向. 我们知道 $\boldsymbol{E}\times\boldsymbol{H}$ 代表电磁场能量流动方向,也就是说,在各向异性媒质中电磁动量与能量传播方向不一致,因而各向异性媒质中平面电磁波的传播方向变得不完全确定. $\boldsymbol{k}$ 的方向代表相速度方向,而 $\boldsymbol{E}\times\boldsymbol{H}$ 方向代表能量流动方向或群速度方向,又称光射线方向.

一个平面波进入单轴媒质以后,总可分解成两个线偏振波的叠加. 其中一个的电场矢量与 $\boldsymbol{k}$ 和光轴所构成的平面垂直,构成寻常波;而另一个电场矢量在 $\boldsymbol{k}$

与光轴构成的平面以内，构成非寻常波．它们以不同的相速度传播，这就是各向异性媒质中的双折射(bi-refracting)现象．

§3.2 光波导基本理论

利用光传递信息的历史很久远，然而从近代科技的发展来看，是从贝尔在1880年的初步探索开始的，这一研究的主要困难在于没有合适的光源和传输介质．随着激光器的诞生以及高锟等人提出的利用包层材料的石英玻璃光学纤维以降低损耗这一卓越贡献，光在通信中的应用便有了基础，而光纤就是最初的光波导[14,15]．然而，集成光学所注重的光波导往往是平面薄膜所构成的平板波导和条形波导，这一方面固然是因为平板波导和条形波导几何形状简单，其导模和辐射模可以用简单的数学公式来描述；另一方面是由于平板波导和条形波导是最常用的最基本的介质光波导，并且易于集成[16~19]．

像任何一种导波结构一样，研究光波导也包括下列基本问题：光场沿光波导横截面的分布规律；光场沿光波导纵向的传播规律；对于每一个导波模式而言，以上两项一一对应；光波导受到扰动时，模式间的耦合；信号沿光波导传播时的衰减；信号沿光波导传播时的畸变；光纤中的非线性效应；光场偏振态沿光波导的演化规律；有源光纤；光波导的激励．在此基础上还面临“综合”问题，即如何设计光波导或相关器件，使之满足给定性能．

现在基本的研究方法有几何光学与波动光学方法．当光波中存在一个或少数几个导波模时，传播常数的分立特性表现得极为明显，用几何光学方法将导致极大误差．换而言之，必须用波动光学方法处理单模(或少模)光波导．同时由于几何光学方法完全忽略了约束光在全反射面以外的存在，不能用以直接处理诸如包层材料引起的损耗、光波导之间的能量耦合以及光纤中稳态分布的建立过程等各种与全反射面以外的电磁场有关的问题．这时必须应用波动光学方法，或在多模光波导情况下对几何光学方法加以修正．

3.2.1 平板波导理论

最简单的平板波导如图3.5所示，它由三层材料构成，中间一层是折射率为n_1的波导薄膜，它沉积在折射率为n_2的基底上，薄膜上面是折射率为n_3的覆盖层，一般都为空气．薄膜的厚度一般在微米量级，可与光的波长相比较．波导薄膜的折射率必须大于基底和覆盖层的折射率，这样光能限制在薄膜之中传播．在下面的讨论中，我们假定导波光是相干单色光，并假定光波导由无损耗、各向同性、非磁性的无源介质构成．

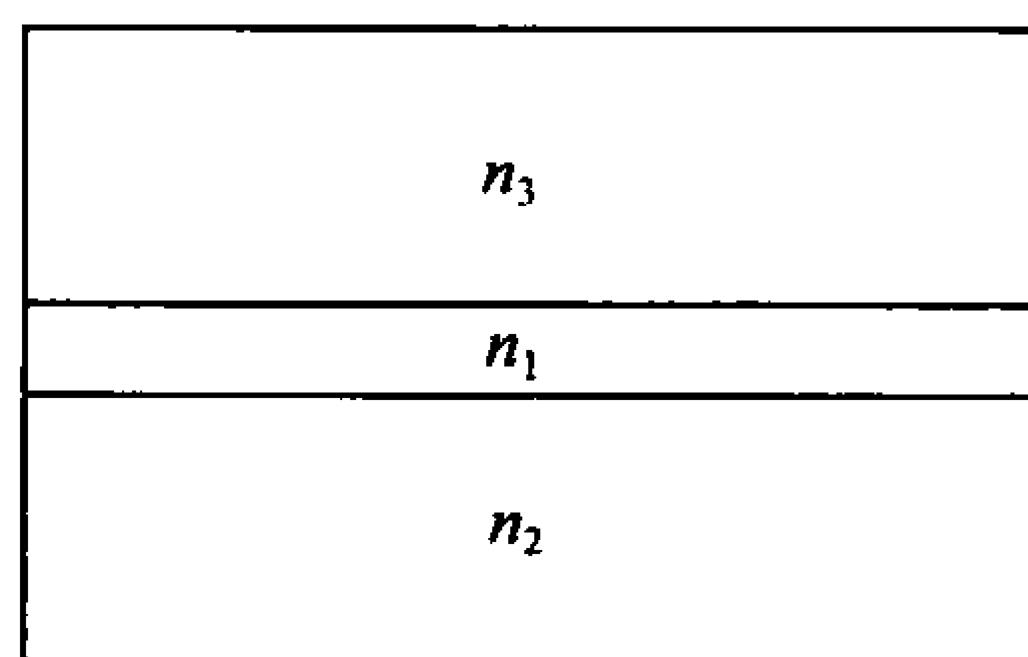


图 3.5 平板介质波导截面图

1. 平板波导的射线光学模型

光在平板波导中的传播可以看做是光线在薄膜-基底和薄膜-覆盖层分界面上的全反射，在薄膜中沿 Z 字形路线传播(参见图 3.6).

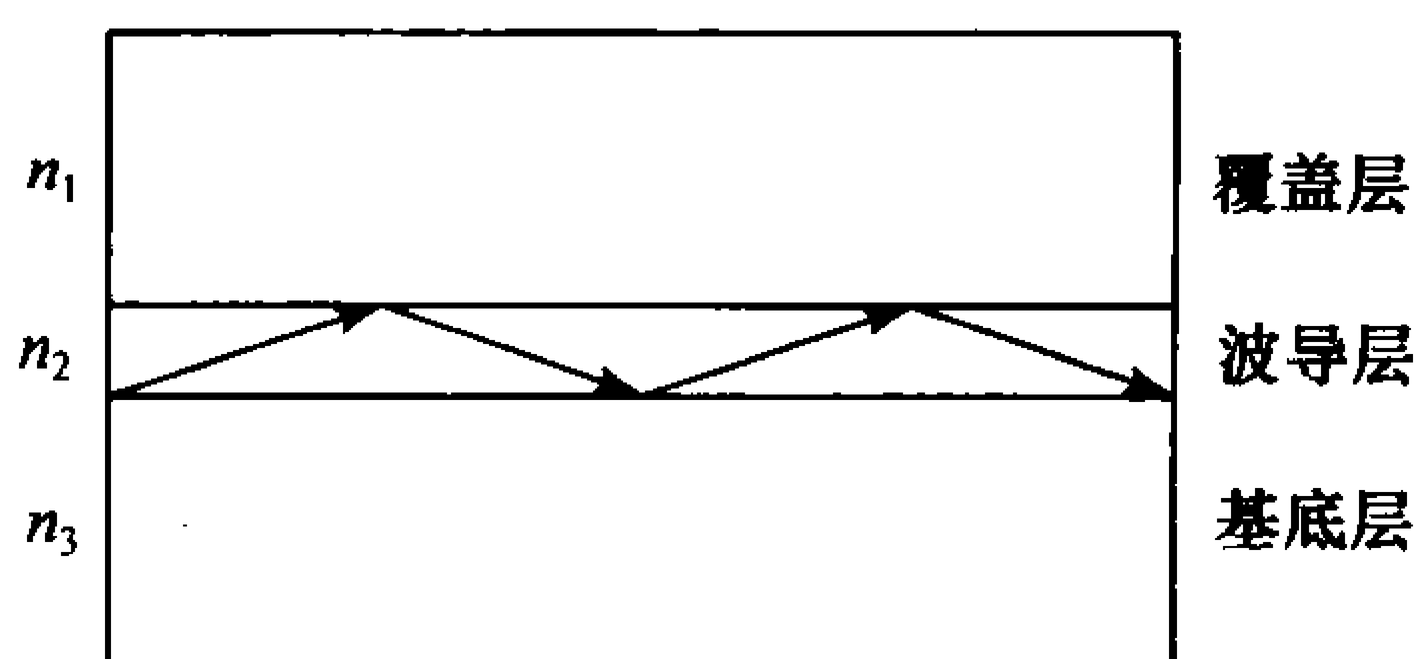


图 3.6 平板波导光传播示意图

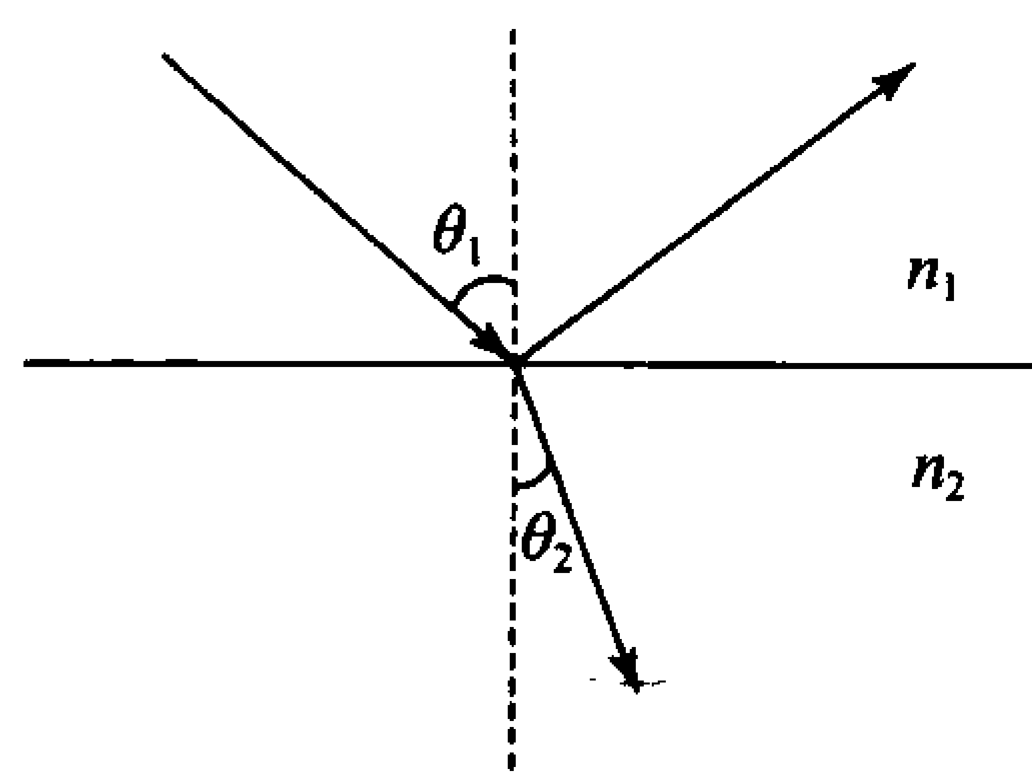


图 3.7 光的反射与折射

我们知道，当光在两介质表面发生折射时(参见图 3.7)，根据折射定律，折射光的出射角应满足

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{3.2.1}$$

在平板波导中，$n_1 > n_2$ 且 $n_1 > n_3$，当入射光的入射角 θ_i 超过临界角 θ_0 时，

$$\sin\theta_0 = \frac{n_1}{n_2}, \tag{3.2.2}$$

入射光发生全发射，此时，在反射点发生一定的相位跃变. 我们从菲涅耳衍射公式出发，推导出发射点的位相跃变 ϕ_{TE}，ϕ_{TM} 为

$$\tan\phi_{TE} = \frac{\sqrt{\beta^2 - k_0^2 n_1^2}}{\sqrt{k_0^2 n_1^2 - \beta^2}}, \tag{3.2.3}$$

$$\tan\phi_{TM} = \frac{n_1^2 \sqrt{\beta^2 - k_0^2 n_1^2}}{n_2^2 \sqrt{k_0^2 n_1^2 - \beta^2}}, \tag{3.2.4}$$

式中 $\beta = k_0 n_1 \sin\theta_i$ 为光的传播常数，$k_0 = 2\pi/\lambda$ 为光在真空中的波数，λ 是光的波长.

我们再来考虑光在光波导中传播. 光在光波导中形成导模，也就是光能全部限制在光波导中进行传播，这就要求光在薄膜-基底和薄膜-覆盖层界面上发生全反射.

光在波导中以锯齿形沿 z 方向传播(如图 3.8 所示),光在 x 方向收到约束,而在 y 方向不受约束.要使光在波导中稳定的传播,就要求

$$2kh - \phi_{12} - \phi_{13} = 2m\pi, \quad m = 0,1,2,\cdots, \tag{3.2.5}$$

其中 $k=k_0 n_1 \cos\theta$,且 ϕ_{12},ϕ_{13} 为全反射的相位差,h 为波导的厚度,m 为模序数,即从零开始取正整数.所以,只有入射角满足式(3.2.5)的光才能在光波导中稳定地传播,我们把式(3.2.5)叫做平板波导的色散方程.

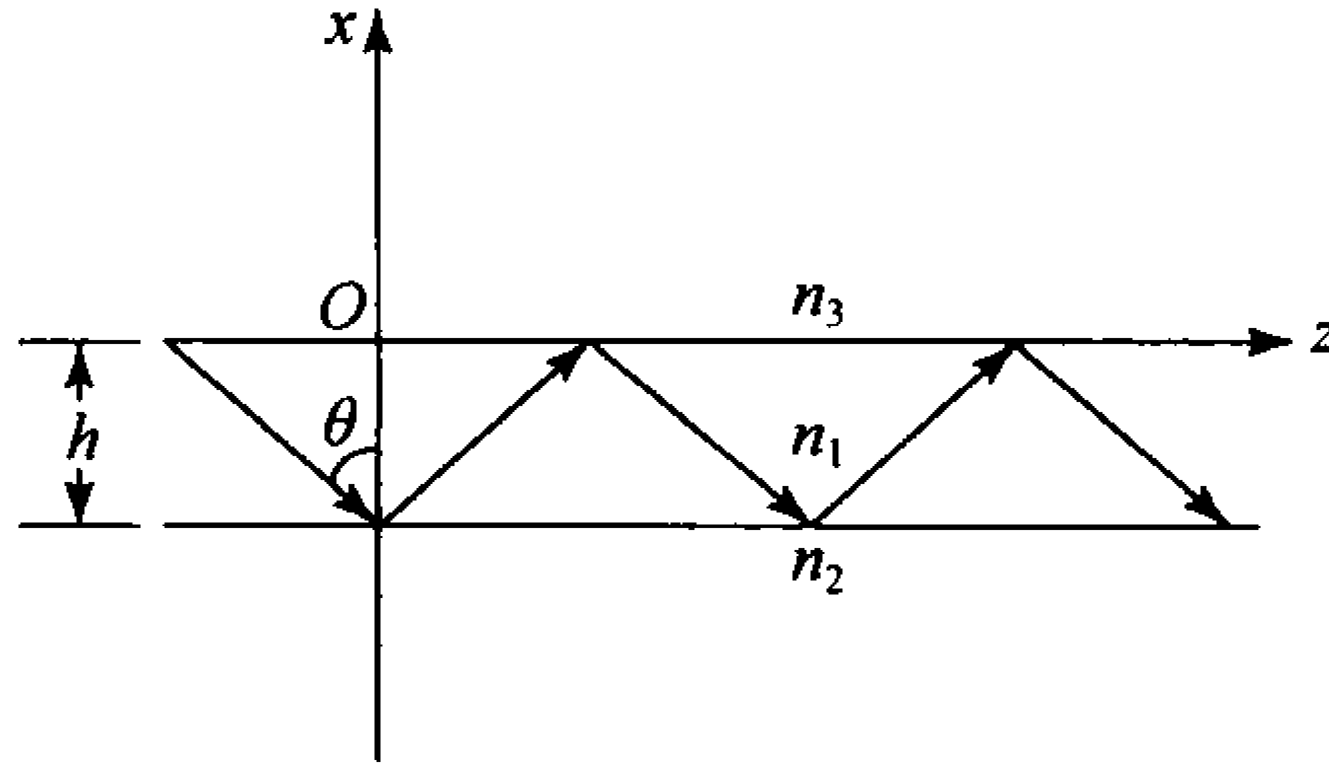


图 3.8 光在波导中的传播

将式(3.2.3)和(3.2.4)分别代入式(3.2.5),可得到与两种偏振态有关的平板波导模式本征方程.对 TE 模,有

$$kh = m\pi + \arctan\left(\frac{p}{k}\right) + \arctan\left(\frac{q}{k}\right), \quad m = 0,1,2,\cdots, \tag{3.2.6}$$

式中

$$k = (k_0^2 n_1^2 - \beta^2)^{1/2}, \tag{3.2.7}$$

$$p = (\beta^2 - k_0^2 n_2^2)^{1/2}, \tag{3.2.8}$$

$$q = (\beta^2 - k_0^2 n_3^2)^{1/2}; \tag{3.2.9}$$

对 TM 模,有

$$kh = m\pi + \arctan\left(\frac{n_1^2 p}{n_2^2 k}\right) + \arctan\left(\frac{n_1^2 q}{n_3^2 k}\right), \tag{3.2.10}$$

其中 β 为光的传播常数,从上式可以看出,β 介于平面光波程基底和薄膜的波数之间,即有

$$k_0 n_2 \leqslant \beta \leqslant k_0 n_1. \tag{3.2.11}$$

为了方便,我们定义波导的有效折射率

$$N = \frac{\beta}{k_0} = n_1 \sin\theta. \tag{3.2.12}$$

N 又称为模折射率,它的取值范围为 $n_2 \leqslant N \leqslant n_1$,利用有效折射率,可将平板波导的模式本征方程改写为

$$(n_1^2 - N^2)^{1/2} k_0 h = m\pi + \arctan\left(\frac{N^2 - n_2^2}{n_1^2 - N^2}\right)$$

$$+\arctan\left(\frac{N^2-n_3^2}{n_1^2-N^2}\right) \quad (\text{TE 模}), \tag{3.2.13}$$

$$(n_1^2-N^2)^{1/2}k_0h = m\pi+\arctan\left(\frac{n_1^2}{n_2^2}\right)\left(\frac{N^2-n_2^2}{n_1^2-N^2}\right)$$

$$+\arctan\left(\frac{n_1^2}{n_3^2}\right)\left(\frac{N^2-n_3^2}{n_1^2-N^2}\right) \quad (\text{TM 模}). \tag{3.2.14}$$

2. 平板波导的电磁理论

平板波导结构中光波的麦克斯韦方程为

$$\nabla\times \boldsymbol{E}(\boldsymbol{r},t) = -\frac{\partial \boldsymbol{B}(\boldsymbol{r},t)}{\partial t}, \tag{3.2.15}$$

$$\nabla\times \boldsymbol{H}(\boldsymbol{r},t) = -\frac{\partial \boldsymbol{D}(\boldsymbol{r},t)}{\partial t}. \tag{3.2.16}$$

$\boldsymbol{E},\boldsymbol{H},\boldsymbol{D}$ 和 $\boldsymbol{B}$ 为随时间和空间变化的电场、磁场、电位移和磁感应强度. 只考虑单色光波且光在平板波导中 y 方向不受限制,沿 z 方向传播,有

$$\begin{aligned}\boldsymbol{E}(x,z,t) &= \boldsymbol{E}(x)\exp[\mathrm{i}(\beta z-\omega t)],\\ \boldsymbol{H}(x,z,t) &= \boldsymbol{H}(x)\exp[\mathrm{i}(\beta z-\omega t)],\end{aligned} \tag{3.2.17}$$

其中 β 为沿 z 方向的传播函数,将式(3.2.17)代入式(3.2.16),(3.2.15),得

$$\begin{cases}\beta E_y = -\mu\omega H_x,\\ \mathrm{i}\beta H_x - \dfrac{\partial H_z}{\partial x} = -\mathrm{i}\omega E_y,\\ \dfrac{\partial E_y}{\partial x} = \mathrm{i}\omega\mu H_z.\end{cases} \tag{3.2.18}$$

$$\begin{cases}\beta H_y = \omega\varepsilon E_x,\\ \dfrac{\partial H_y}{\partial x} = -\mathrm{i}\omega\varepsilon E_z,\\ \mathrm{i}\beta E_x - \dfrac{\partial E_z}{\partial x} = \mathrm{i}\omega\mu H_y.\end{cases} \tag{3.2.19}$$

由上述 6 个方程可以看出,麦克斯韦方程变为两组独立的方程,一组只含 E_y,H_x,H_z,而另一组方程只含 E_x,E_z,H_y. 前者称为 TE 波,即电场垂直于波的入射面;后者称为 TM 波,即磁场垂直于波的入射面. 化简式(3.2.18),得

$$\frac{\partial^2 E_y}{\partial x^2}+(k_0^2n_j^2-\beta^2)E_y = 0; \tag{3.2.20}$$

化简式(3.2.19),得

$$\frac{\partial^2 H_y}{\partial x^2}+(k_0^2n_j^2-\beta^2)H_y = 0, \tag{3.2.21}$$

式中 $k_0=\omega\sqrt{\varepsilon_0\mu_0}=2\pi/\lambda$,是光在真空中的传播常数. 式(3.2.20)和(3.2.21)分别称为 TE 波和 TM 波的亥姆霍兹方程.

考虑到波导无源、无损耗、各向同性和非磁性，式(3.2.20)和(3.2.21)应满足边界条件

$$\begin{cases} \boldsymbol{e}_n \cdot (\boldsymbol{D}_1 - \boldsymbol{D}_2) = 0, \\ \boldsymbol{e}_n \cdot (\boldsymbol{B}_1 - \boldsymbol{B}_2) = 0, \\ \boldsymbol{e}_n \cdot (\boldsymbol{D}_1 - \boldsymbol{D}_3) = 0, \\ \boldsymbol{e}_n \cdot (\boldsymbol{B}_1 - \boldsymbol{B}_3) = 0, \\ \boldsymbol{e}_n \times (\boldsymbol{H}_1 - \boldsymbol{H}_2) = \boldsymbol{0}, \\ \boldsymbol{e}_n \times (\boldsymbol{E}_1 - \boldsymbol{E}_2) = \boldsymbol{0}, \\ \boldsymbol{e}_n \times (\boldsymbol{E}_1 - \boldsymbol{E}_3) = \boldsymbol{0}, \\ \boldsymbol{e}_n \times (\boldsymbol{H}_1 - \boldsymbol{H}_3) = \boldsymbol{0}. \end{cases} \tag{3.2.22}$$

对 TE 模，由于我们分析的是导模，从而可知在薄膜内是振荡场，即是驻波. 因而，三层平板波导的电场分布为

$$E_y(x) = \begin{cases} A\exp(-qx), & 0 \leqslant x \leqslant \infty, \\ B\cos(kx) + C\sin(kx), & -h \leqslant x \leqslant 0, \\ A\exp[p(x+h)], & -\infty \leqslant x \leqslant -h, \end{cases} \tag{3.2.23}$$

式中 A,B,C,D 是边界条件所确定的常数，参数 k,p,q 由式(3.2.7)，(3.2.8)，(3.2.9)所定义. 考虑边界条件，即 E_y 在薄膜-覆盖层和薄膜-基底界处连续及 $\frac{\partial E_y}{\partial x}$ 在薄膜-覆盖层界出连续，可得

$$\begin{cases} A = B, \\ B\cos(kh) - C\sin(kh) = A, \\ -qA = kC. \end{cases} \tag{3.2.24}$$

同时，利用归一化条件，即消去待定常数. 利用 $\frac{\partial E_y}{\partial x}$ 在薄膜-基底处连续，可得到平板波 TE 模的模式本征方程

$$\tan(kh) = \frac{p+q}{k\left(1-\frac{pq}{k^2}\right)}. \tag{3.2.25}$$

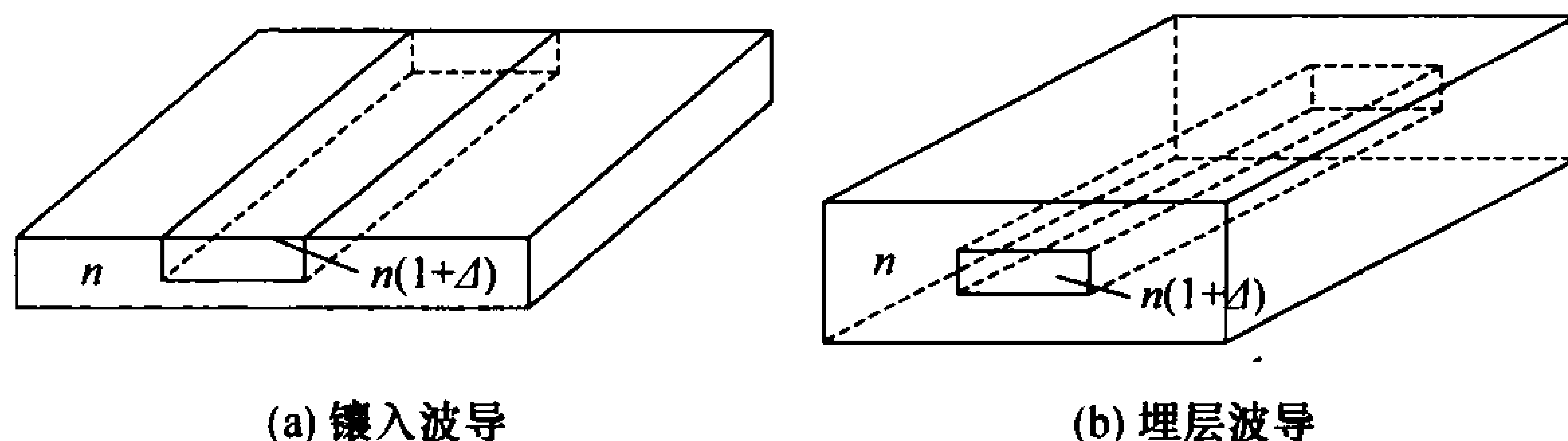
通过求解式(3.2.25)可解出模式本征值. 可以看出，式(3.2.25)与(3.2.6)完全一致. 同样，可解出 TM 模的模式本征方程

$$\tan(kh) = \frac{n_1^2 k(n_3^2 p + n_2^2 q)}{n_2^2 n_3^2 k - n_1^4 pq}. \tag{3.2.26}$$

3.2.2 矩形波导理论

介质平板波导的电磁场仅在一个方向受到限制[20~22]. 光场在另一个方向上不受限制. 光场在介质平板波导中传播时要沿这个非束缚方向发散. 因而，在实

际的集成光路中，经常使用的是能在横截面的二维方向上限制光场能量的条形介质波导. 其结构如图 3.9(a)和(b)所示，条形介质波导分为镶入波导和埋层波导.



(a) 镶入波导 (b) 埋层波导

图 3.9 条形波导结构示意图

对矩形介质波导的研究，常常把条形波导的截面笼统地作为矩形波导来处理. 即使这样，还牵涉到复杂的二维电磁场问题，要得到严格的解析解几乎是不可能的. 故而采用数值计算方法，可求得近似解，并且可以达到所希望的精确度. 下面，我们来讨论一下马卡提里近似解析法和有效射率法.

1. 马卡提里近似解析法

波导结构如图 3.10 所示，折射率为 n_1 的波导区被折射率为 n_2, n_3, n_4, n_5 的介质层包围.

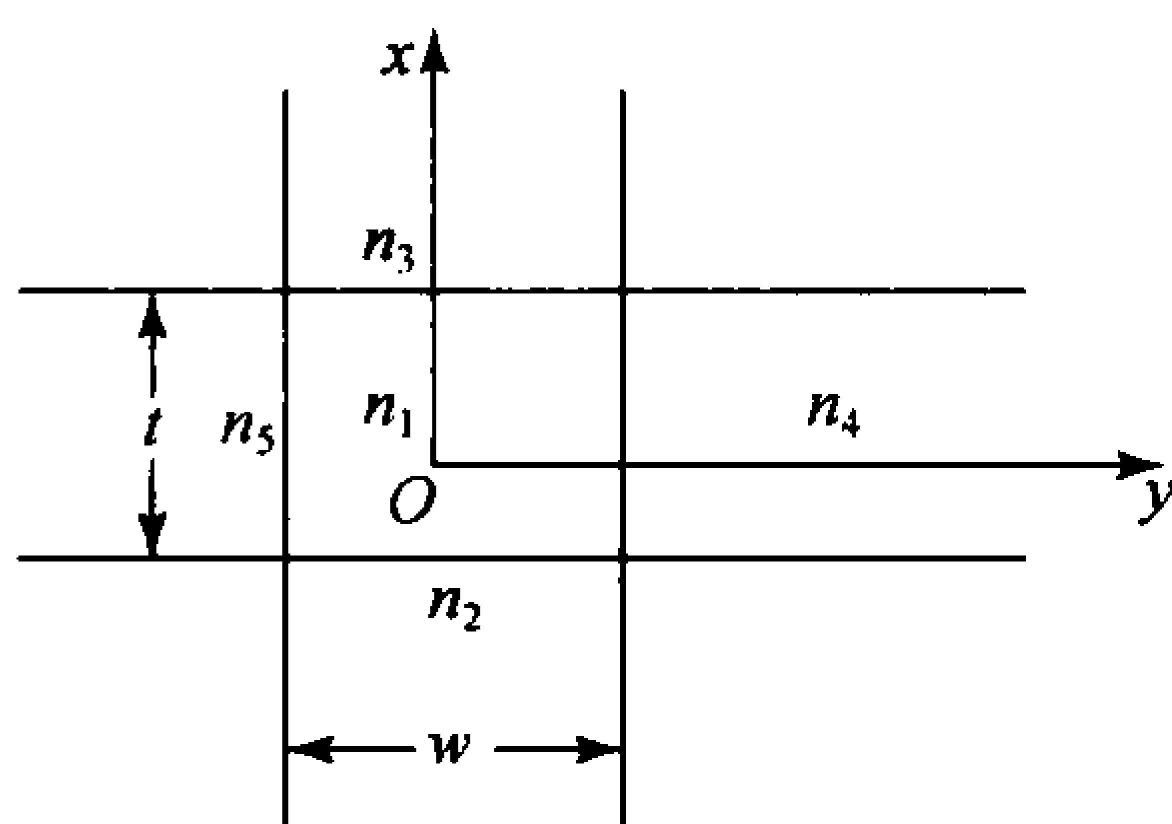


图 3.10 条形波导结构示意图

在每个区域中，波数用 k_j 表示：

$$k_{jx}^2 + k_{jy}^2 + k_{jz}^2 = k_j^2, \quad j = 1,2,3,4,5. \tag{3.2.27}$$

从图 3.10 中可以看出，这里有一个非常复杂的电磁场边值问题. 我们假定两条：① 介质波导中的传播模式处于远离截止模式. 光能在波导芯区中传播，参照 3.2.1 小节中我们讨论过的平板波导，不难得到

$$\begin{cases} k_{1y} = k_{4y} = k_{5y} = k_y, \\ k_{1x} = k_{4x} = k_{5x} = k_x. \end{cases} \tag{3.2.28}$$

② 假定包层的折射率 n_2, n_3, n_4, n_5 比 n_1 略小一些，即弱导条件

$$\frac{n_1}{n_i}-1 \ll 1, \quad i=2,3,4,5, \tag{3.2.29}$$

这样，可将问题大大简化.

下面，我们从电磁场理论出发，采用类似于平板波导的方法，先设电磁场随时间 t 和空间轴 z 的变化仍为 $\exp[\mathrm{i}(\beta z-\omega t)]$，对导模有

$$k_{1z}=k_{2z}=k_{3z}=k_{4z}=k_{5z}=\beta_0. \tag{3.2.30}$$

根据麦克斯韦方程可以导出，矩形波导的波动方程为

$$\frac{\partial^2\varphi}{\partial x^2}+\frac{\partial^2\varphi}{\partial y^2}+(k_0 n_j^2-\beta^2)\varphi=0. \tag{3.2.31}$$

对 E_{mn}^x 模式来说，E_x 和 H_y 是主要横场分量，故有

$$H_y=\begin{cases} H_1\cos(k_x x+\xi)\cos(k_y y+\eta), \\ H_2\cos(k_x y+\eta)\exp(k'_{2x}x), \\ H_3\cos(k_x y+\eta)\exp(-k'_{3x}x), \\ H_4\cos(k_x x+\xi)\exp(-k'_{4x}y), \\ H_5\cos(k_x x+\xi)\exp(k'_{5y}y), \end{cases} \tag{3.2.32}$$

式中 H_1,H_2,H_3,H_4,H_5 均为常数振幅因子，其下标表示各个区域的代号，ξ,η 为任意位相因子，其中

$$\begin{cases} k'^2_{2x}=-(k_0^2n_2^2-k_y^2-\beta^2)=-k_{2x}^2, \\ k'^2_{3x}=-(k_0^2n_3^2-k_y^2-\beta^2)=-k_{3x}^2, \\ k'^2_{4y}=-(k_0^2n_4^2-k_x^2-\beta^2)=-k_{4x}^2, \\ k'^2_{5y}=-(k_0^2n_5^2-k_x^2-\beta^2)=-k_{5x}^2. \end{cases} \tag{3.2.33}$$

设 $H_x=0$，根据近似条件(3.2.29)，并代入边界条件，化简方程，得

$$k_0t=m\pi+\arctan\left(\frac{n_1^2k'_{2x}}{n_2^2k_x}\right)+\arctan\left(\frac{n_1^2k'_{3x}}{n_3^2k_x}\right), \quad m=0,1,2,\cdots. \tag{3.2.34}$$

利用三角函数的周期性，可以写成

$$k_0t=m\pi+\arctan\left(\frac{n_2^2k'_x}{n_1^2k'_{2x}}\right)+\arctan\left(\frac{n_3^2k'_x}{n_1^2k'_{3x}}\right), \quad m=0,1,2,\cdots. \tag{3.2.35}$$

由式(3.2.27)，(3.2.28)，(3.2.29)可知

$$\beta^2=k_0^2n_1^2-(k_x^2+k_y^2). \tag{3.2.36}$$

由式(3.2.33)，可得

$$k'^2_{2x}=k_0^2(n_1^2-n_2^2)-k_x^2, \tag{3.2.37}$$

$$k'^2_{3x}=k_0^2(n_1^2-n_3^2)-k_x^2. \tag{3.2.38}$$

在 $y=\pm\omega/2$，H_y 和 H_z 在边界上连续，故可得方程

$$k_y\omega=n\pi-\arctan\left(\frac{k_y}{k'_{4y}}\right)-\arctan\left(\frac{k_y}{k'_{5y}}\right), \quad n=1,2,3,\cdots, \tag{3.2.39}$$

式中

$$k_{4y}'^2 = k_0^2(n_1^2 - n_4^2) - k_y^2, \tag{3.2.40}$$

$$k_{5y}'^2 = k_0^2(n_1^2 - n_5^2) - k_y^2. \tag{3.2.41}$$

从上式可看出，在二维矩形介质波导中，必须由两个方程(3.2.35)，(3.2.41)联立求解，分别求出 k_x，k_y，再应用式(3.2.36)，才能求出模式本征值 β.

对 E_{mn}^x 模，采用同样方法，可推出

$$k_x t = m\pi - \arctan\left(\frac{k_x}{k'_{2x}}\right) - \arctan\left(\frac{k_x}{k'_{3x}}\right), \quad n = 1,2,3,\cdots, \tag{3.2.42}$$

$$k_y \omega = n\pi - \arctan\left(\frac{n_4^2 k_y}{n_1^2 k'_{4y}}\right) + \arctan\left(\frac{n_5^2 k_y}{n_1^2 k'_{5y}}\right), \quad n = 1,2,3,\cdots, \tag{3.2.43}$$

从模式本征方程可以看出，它实际上是由两个独立的介质平板波导 TE 与 TM 模的模式本征方程组成.

2. 有效折射率法

在马卡提里近似的基础上，我们来介绍一种更为实用的近似方法——有效折射率法(图 3.11).

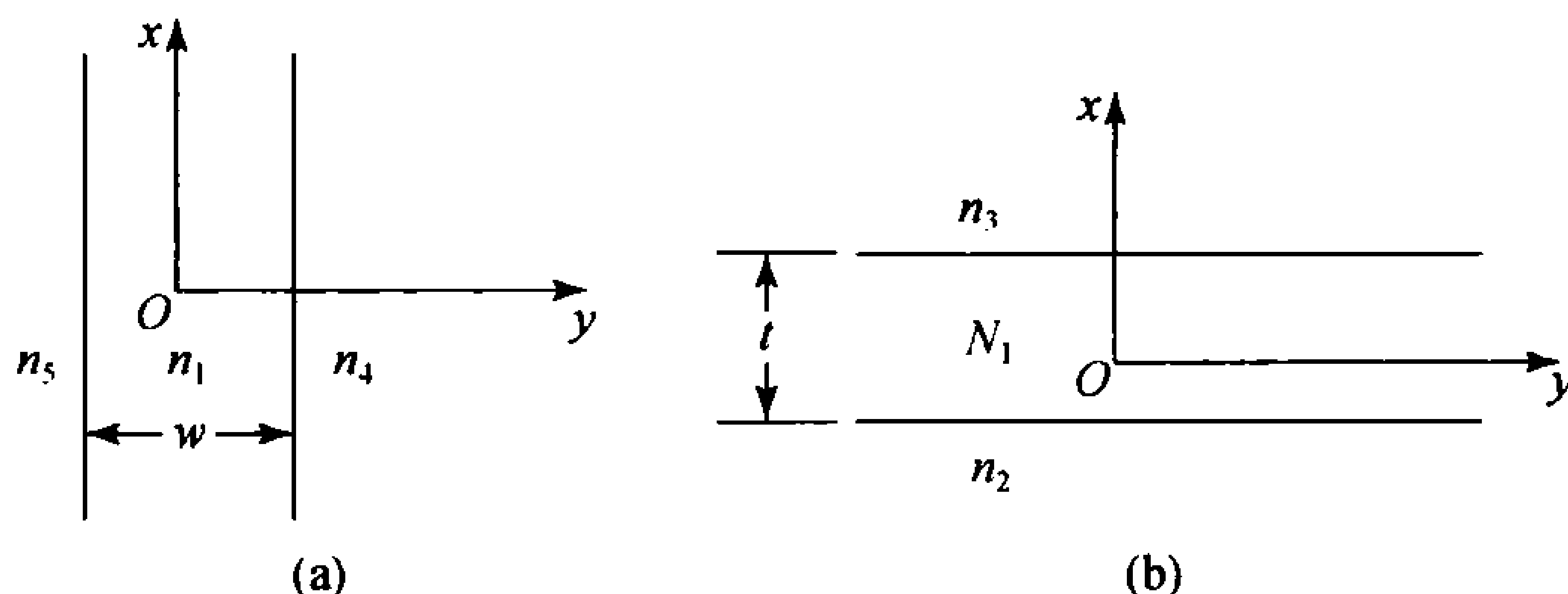


图 3.11　有效折射率示意图

我们把矩形介质波导看做如图 3.11 所示的两个一维介质波导的组合，但这两个不完全独立，其中图 3.11(b)所示的平板波导的导波层的折射率等于图 3.11(a)所示的平板波导的模的有效折射率 N_1：

$$N_1^2 = n_1^2 - \left(\frac{k_y}{k_0}\right)^2. \tag{3.2.44}$$

对 E_{mn}^x 模，已知主要电磁场分量 E_x 和 H_y. 从图 3.11 中可见，对于(a)所示平板波导，这种场相当于 TE 波；对(b)所示平板波导，这种场相当于 TM 波. 故两平板波导的模式本征方程分别为

$$k_y \omega = n\pi - \arctan\left(\frac{n_4^2}{k'_{4y}}\right) - \arctan\left(\frac{k_y}{k'_{5y}}\right), \quad n = 1,2,3,\cdots, \tag{3.2.45}$$

$$k_x t = m\pi - \arctan\left(\frac{n_2^2 k_x}{n_1^2 k''_{2x}}\right) - \arctan\left(\frac{n_3^2 k_x}{n_1^2 k''_{3x}}\right), \quad m = 1,2,3,\cdots, \tag{3.2.46}$$

式中

$$k''^2_{2x} = k_0^2(N_1^2 - n_2^2) - k_x^2, \tag{3.2.47}$$

$$k''^2_{3x} = k_0^2(N_1^2 - n_3^2) - k_x^2. \tag{3.2.48}$$

由式(3.2.46)可知,矩形波导的传播常数为

$$\beta = (k_0^2 N_1^2 - k_x^2)^{1/2}. \tag{3.2.49}$$

对 E^y_{mn} 模式来说,已知主要电磁场分量是 E_y 和 H_x. 从图 3.11 得知,这种场相当于图 3.11(a)所示的平板 TM 波,对图 3.11(b)所示的平板波导,相当于 TE 波,故而模式本征方程为

$$k_y\omega = m\pi - \arctan\left(\frac{n_4^2 k_y}{n_4^2 k_{4y}}\right) - \arctan\left(\frac{n_5^2 k_y}{n_4^2 k'_{5y}}\right), \quad m = 1,2,3,\cdots, \tag{3.2.50}$$

$$k_x t = n\pi - \arctan\left(\frac{k_x}{k'_{2x}}\right) - \arctan\left(\frac{k_x}{k'_{3x}}\right), \quad n = 1,2,3,\cdots. \tag{3.2.51}$$

从上面的模式本征方程可以看出,两个不同的近似方法有相似的解,但式中得 k_z 却不相同.

通过计算比较,两种方法在远离截止区时符合得很好,在接近截止区,有效折射率法比马卡提里近似法更精确.

§3.3 波导耦合理论

光波导能量交换和传输的本质是波导模式之间的耦合,耦合的意义包含能量从一个波导进入另一个波导,空间光波的能量进入波导[23,24]. 本节在弱耦合条件下简述两个结构相同的平行条形通道波导间的横向耦合理论,阐述经典的耦合系数计算方法,并针对经典的耦合系数,提出了改进的计算波导横向耦合系数的公式,并推导了单模条形通道波导间的耦合效率及振幅传递系数公式;简述了波导间纵向耦合的理论模型及光波由空间耦合进入波导的理论模型,并将罗兰(Rowland)圆耦合器模型的空间耦合效率计算方法进一步推广到非傍轴情况,导出了更为客观普遍的空间耦合效率计算方法,目的是为阵列波导光栅的设计提供正确而且严格的耦合理论.

3.3.1 横向弱耦合理论

横向耦合实际上是光场能量由一个波导进入另一个并列波导的光学隧道现象,因为能量以相干的方式传输,使得传输方向不变,可以通过此类效应制作定向耦合器,其结构有多层平面结构或通道波导结构. 但是横向耦合也会在集成光学器件,如阵列波导光栅(arrayed waveguide grating, AWG)中引起波导间的串扰.

1. 平面波导耦合器

尽管可以通过对接的方式将两个波导耦合起来，使光场能量从一个波导进入另一个波导，但更普遍的方法是通过平板波导横向耦合，平板波导横向耦合器的结构如图 3.12 所示.

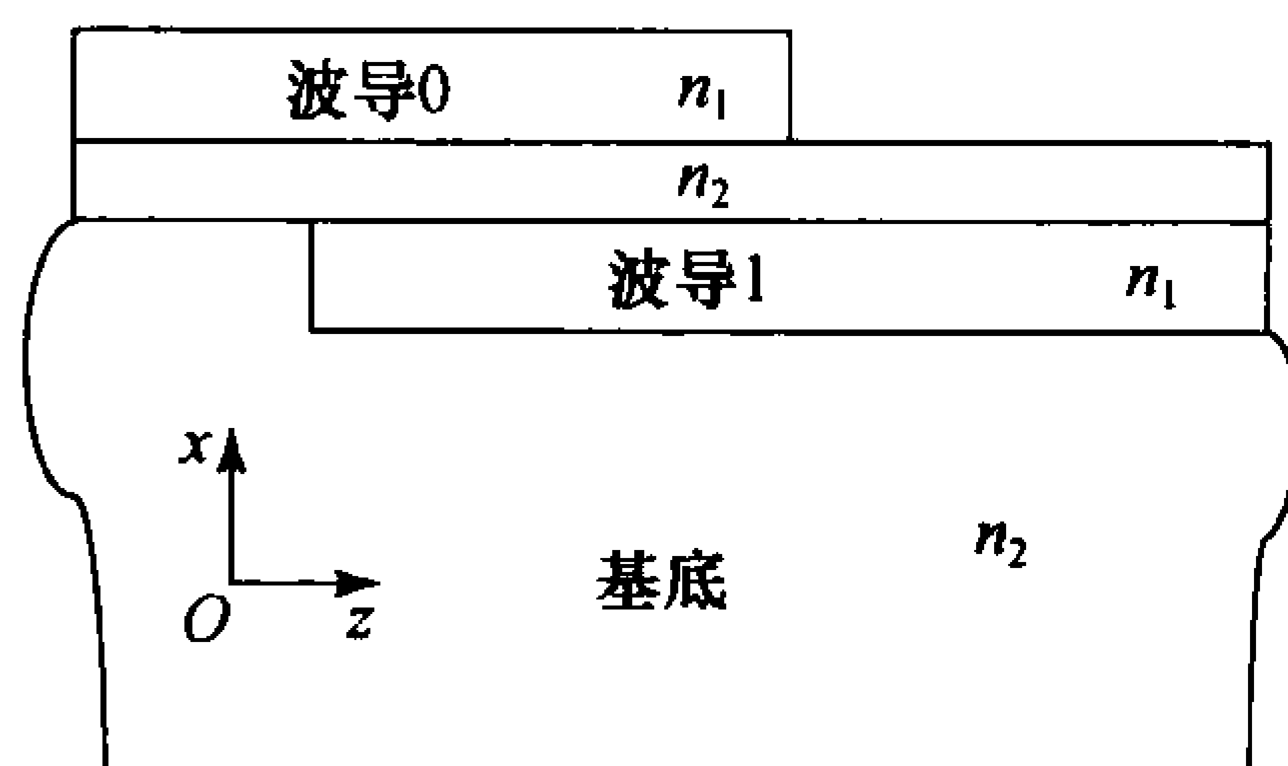


图 3.12　平板波导耦合器

2. 双通道定向波导耦合器

双通道波导耦合器如图 3.13 所示. 对于平板波导耦合器或双通道波导耦合器，耦合都只由波导横方向的模场引起. 在一定的作用长度，能量将会交替地从一个波导进入另一个波导.

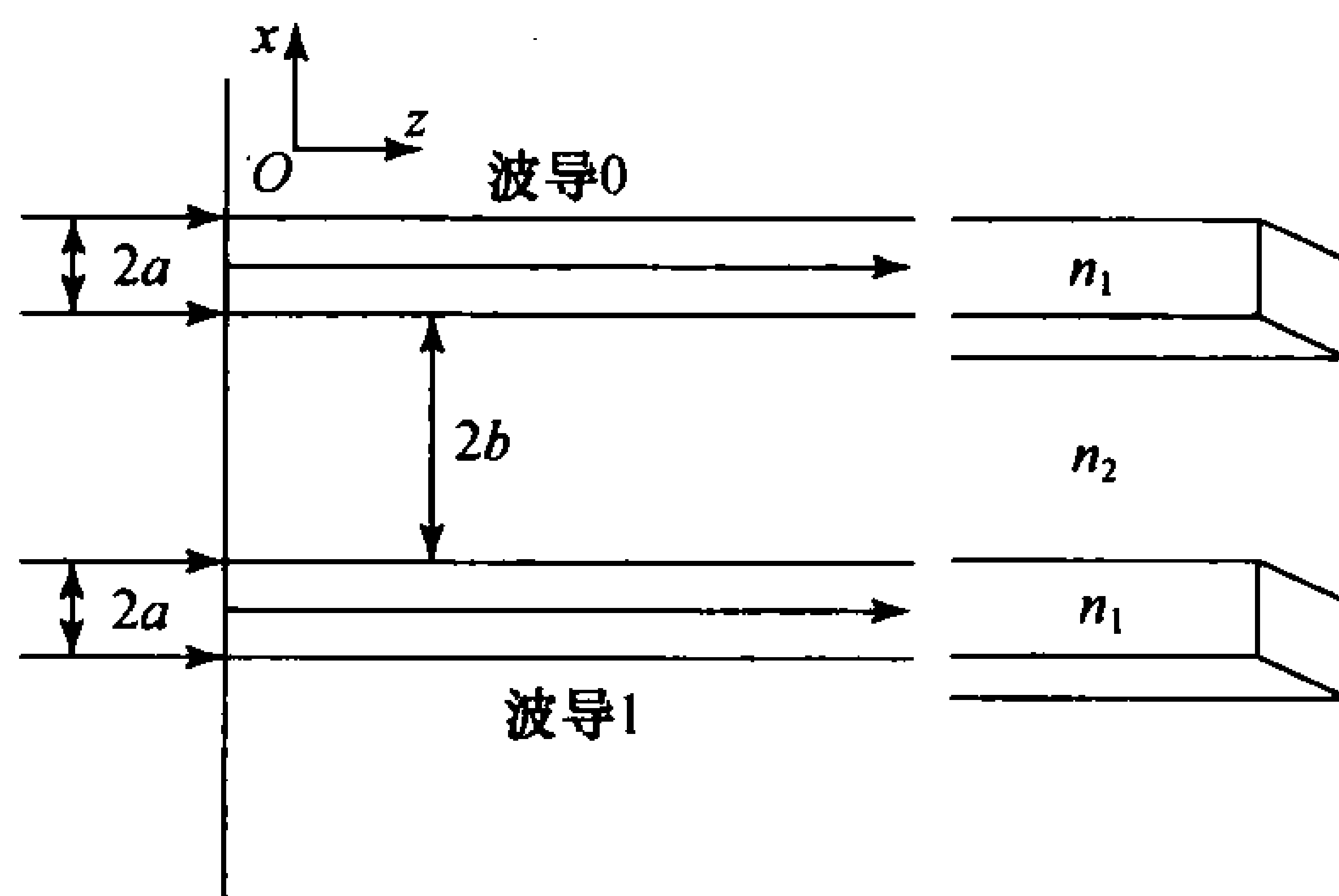


图 3.13　双通道条形波导耦合器

3. 理论模型

当两个波导间距与波导本身线度相比足够大，耦合作用不足以改变两个参与相互耦合作用的波导模式场分布，而只改变振幅系数时，波导间的横向耦合满足弱耦合条件.

可以用 Yariv 的耦合理论方法，来描述同步耦合模的耦合理论.

两个模之间的耦合可以通过两个模振幅之间的普通模耦合方程导出，以 $A_0(z)$，$A_1(z)$来代表图 3.12，图 3.13 中波导 1，2 振幅. 两个模之间满足的关系表述为

$$\frac{\mathrm{d}A_0(z)}{\mathrm{d}z} = -\mathrm{i}\beta_0 A_0(z) + k_{01}A_1(z), \tag{3.3.1}$$

$$\frac{\mathrm{d}A_1(z)}{\mathrm{d}z} = -\mathrm{i}\beta_1 A_1(z) + k_{10}A_0(z), \tag{3.3.2}$$

其中 k_{10},k_{02} 为波导之间相互耦合系数,而 β_1,β_2 为传输参量. 当两个波导结构完全相同时,$k_{10}=k_{02}=\mathrm{i}k$,传输参量也相同,满足

$$\beta_1 = \beta_2 = \beta = \beta_r - \mathrm{i}\,\frac{a}{2}, \tag{3.3.3}$$

其中 a 为指数式的光损耗系数.

假设光在 $z=0$ 处进入波导 0,则方程(3.3.1),(3.3.2)的边界条件为

$$A_0(0) = 1,\quad A_1(0) = 0. \tag{3.3.4}$$

于是方程(3.3.1),(3.3.2)的解为

$$A_0(z) = \cos(kz)\exp(\mathrm{i}\beta z), \tag{3.3.5}$$

$$A_1(z) = \sin(kz)\exp(\mathrm{i}\beta z). \tag{3.3.6}$$

波导内的功率流随传输距离 z 交替变化,分别为

$$P_0(z) = \cos^2(kz)\mathrm{e}^{-az}, \tag{3.3.7}$$

$$P_1(z) = \sin^2(kz)\mathrm{e}^{-az}, \tag{3.3.8}$$

两个相互耦合波导的功率分布曲线如图 3.14 所示.

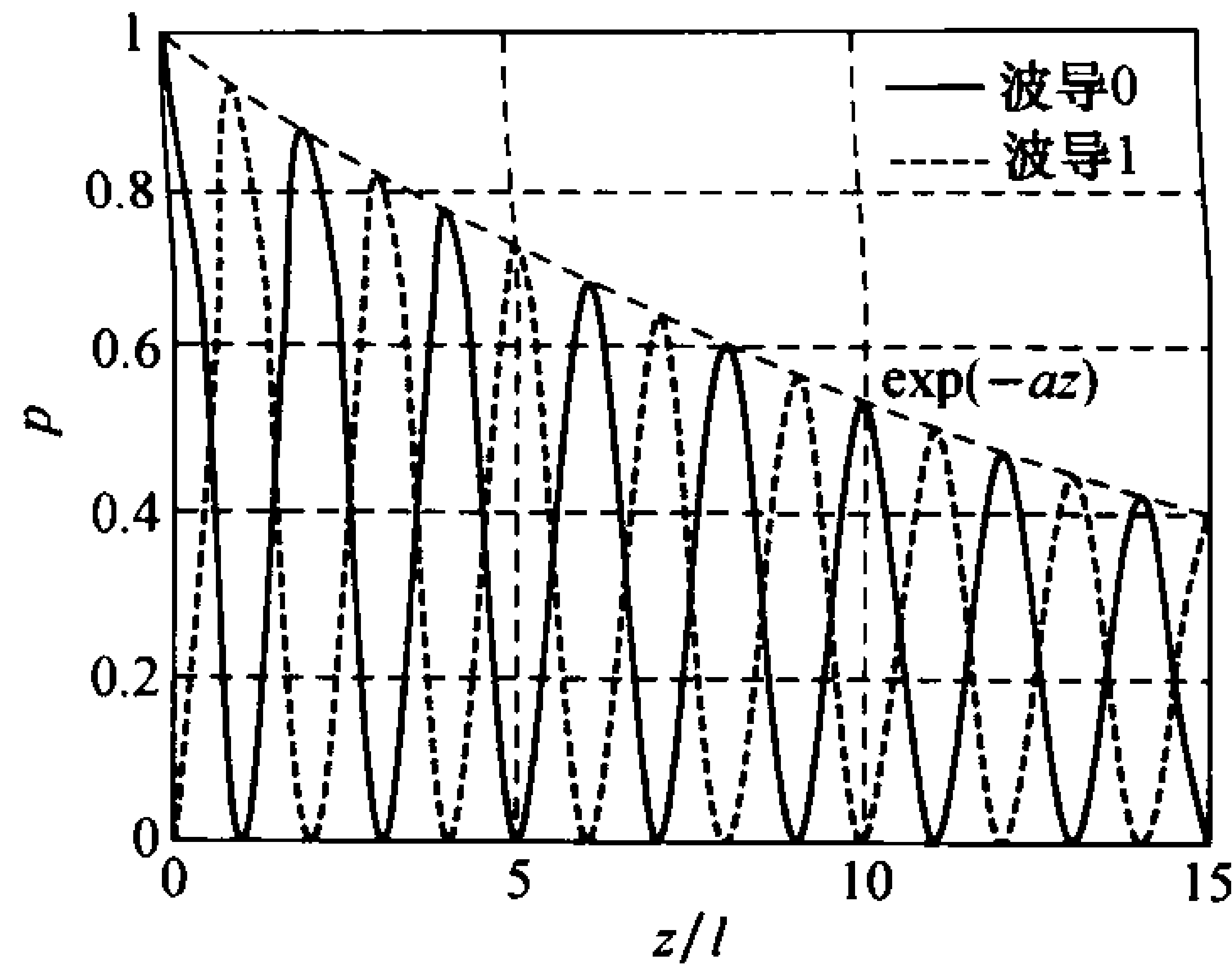


图 3.14 波导定向耦合功率分布,其中 $l=\pi/(2k)$

功率完全由波导 0 转移至波导 1 时,相互作用的耦合长度满足

$$L = \frac{\pi}{2k} + \frac{m\pi}{k}, \tag{3.3.9}$$

其中 $m=1,2,3,\cdots$,可见,对于耦合器来说,其耦合长度 L 是一个至关重要的参量. 由于实际波导总有吸收损耗,传输常数 β 是个复数,因而两个波导的总功率按因子 $\exp(-az)$衰减.

4. 耦合系数

耦合系数的计算表达式为

$$k=\frac{\omega\varepsilon_0}{4P}\int_{-\infty}^{\infty}(n_1^2-n_2^2)\psi(x-2b-2a)\mathrm{d}x,\tag{3.3.10}$$

其中 P 为单个波导的传输功率,以平板波导 TE_0 模为例,根据 TE_0 模场分布函数得

$$P=\frac{\beta\alpha}{2\omega\mu_0}(1+1/W),\tag{3.3.11}$$

其中 W 表示包层衰减参量,$W=a\sqrt{\beta^2-k_0^2n^2}$,$n$ 为包层折射率,a 为芯层半宽度.可以将耦合系数的计算表达式化简为

$$k=\frac{Wv^2}{2a^2\beta(1+W)}\int_{-\infty}^{\infty}\psi(x)\psi(x-2b-2a)\mathrm{d}x.\tag{3.3.12}$$

耦合系数 k 明显依赖于波导中模尾的形状,一般文献中耦合系数忽略了包层光场对耦合产生底作用,积分区间只考虑波导 1 的芯层,即积分区域为[$-a,a$]得到的结果为

$$\begin{aligned}k&=\frac{Wv^2}{2a^2\beta(1+W)}\int_{-\infty}^{\infty}\psi(x)\psi(x-2b-2a)\mathrm{d}x\\&=\exp(-2Wb/a)\frac{Wv^2}{2a^2\beta(1+W)}.\end{aligned}\tag{3.3.13}$$

然而,实际上耦合系数的计算公式满足的基本为弱耦合,为了更为客观的描述实际情况,需要考虑波导包层光场的耦合作用,积分区域应按照规定义在整个区域积分,改进的耦合系数应为

$$\begin{aligned}k&=\frac{Wv^2}{2a^2\beta(1+W)}\int_{-\infty}^{\infty}\psi(x)\psi(x-2b-2a)\mathrm{d}x\\&=\exp(-2Wb/a)\frac{Wv^2}{2a^2\beta(1+W)}\left[\frac{2b}{a}+\frac{4W}{v^2}+\frac{\exp(-2W)}{W}\right].\end{aligned}\tag{3.3.14}$$

3.3.2 不考虑输入角度的纵向耦合理论

纵向耦合通常用于波导之间的对接,如矩形波导和光纤之间的耦合,或不考虑光束输入角度影响的情况下,空间光束与波导的纵向耦合.

纵向耦合效率决定于入射光场 $\psi_1(x,y)$与波导场 $\psi_2(x,y)$的重叠积分,即

$$\eta=\frac{\left|\iint\limits_S\psi_1(x,y)\psi_2(x,y)\mathrm{d}x\mathrm{d}y\right|^2}{\iint\limits_S\psi_1^2(x,y)\mathrm{d}x\mathrm{d}y\iint\limits_S\psi_2^2(x,y)\mathrm{d}x\mathrm{d}y}.\tag{3.3.15}$$

设入射光场 $\psi_1(x,y)$的最大振幅为 A_1,耦合进入波导 2 后,在波导 2 中激励起的

最大振幅为 A_2，考虑到耦合效率的定义为

$$\eta_{1\to2}=\frac{P_2}{P_1}=\frac{A_2^2\iint_S\psi_2^2(x,y)\mathrm{d}x\mathrm{d}y}{A_1^2\iint_S\psi_1^2(x,y)\mathrm{d}x\mathrm{d}y\iint_S\psi_2^2(x,y)\mathrm{d}x\mathrm{d}y},\tag{3.3.16}$$

其中 P_1，P_2 分别为波导 1,2 的传输功率，则可以导出最大振幅传递系数 $T_{1\to2}$ 为

$$T_{1\to2}=\frac{A_2}{A_1}=\frac{\iint_S\psi_1(x,y)\mathrm{d}x\mathrm{d}y\iint_S\psi_2(x,y)\mathrm{d}x\mathrm{d}y}{\iint_S\psi_2^2(x,y)\mathrm{d}x\mathrm{d}y}.\tag{3.3.17}$$

3.3.3 考虑光束传输角度的波导空间耦合理论

经典的罗兰圆形耦合器结构如图 3.15 所示，由发送波导和接收波导组成，发射波导端口位于圆周 1 上，端口法线指向 O_2，接受波导位于圆周 2 上，端口法线指向 O_1. 这种情况下，光束由空间耦合进入波导，此时应该考虑光束的传输角度，这样更为客观且符合实际，而经典的耦合效率公式忽略了光束的传输方向.

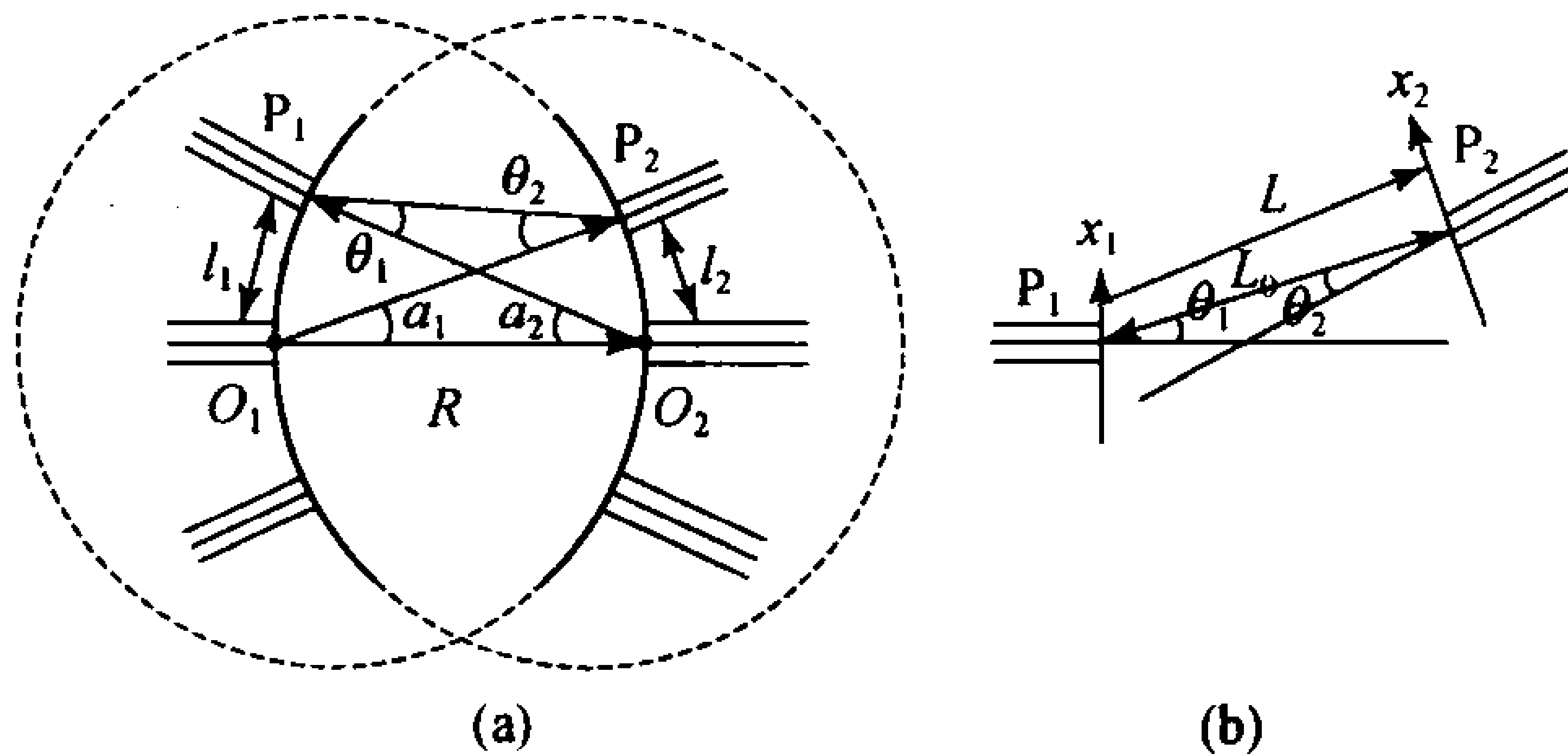


图 3.15 星形结构耦合器

考虑一般的情况，光束由波导 $\mathrm{P_1}$ 发射，$\mathrm{P_2}$ 接收，对于实际星形结构耦合器来说，$R\gg l_1$，$R\gg l_2$，因此

$$a_1\approx\theta_1,\quad a_2\approx\theta_2.\tag{3.3.18}$$

根据瑞利判据[10]，图 3.15(b)中的 L 可以简化为

$$L=L_0-x_1\sin\theta_2-x_2\sin\theta_2,\tag{3.3.19}$$

其中 x_1，x_2 分别为以波导 $\mathrm{P_1}$，$\mathrm{P_2}$ 中心为原点的坐标.

根据瑞利-索末菲衍射积分公式，波导 $\mathrm{P_1}$ 衍射光束在波导 $\mathrm{P_2}$ 入口处产生的光场分布为

$$\psi_1(x_2)=\frac{\cos\theta_1\exp(\mathrm{i}kL_0)}{\sqrt{\mathrm{i}kL_0}}\exp(-\mathrm{i}kx_2\sin\theta_2)\int_{-\infty}^{\infty}\psi_1(x_1)\exp(-\mathrm{i}kx_1\sin\theta_1)\mathrm{d}x_1$$

$$= \frac{a_1 \cos\theta_1 \exp(\mathrm{i}kL_0)}{\sqrt{\mathrm{i}\lambda L_0}} \exp(-\mathrm{i}kx_2 \sin\theta_2) S(\theta_1), \tag{3.3.20}$$

其中 $S(\theta_1)$为波导 P_1 的衍射角谱.

结合式(3.3.20),式(3.3.15)描述的耦合效率可改写为

$$\eta_{1\to 2} = \frac{a_1 a_2 \cos^2\theta_1 \cos^2\theta_2}{\lambda L_0} \frac{S(\theta_1) S(\theta_2)}{\int_{-\infty}^{\infty} \psi_1^2(x_1)\mathrm{d}x_1 \int_{-\infty}^{\infty} \psi_2^2(x_2)\mathrm{d}x_2}, \tag{3.3.21}$$

对应由波导 P_1 到 P_2的最大振幅传递系数为

$$T_{1\to 2} = \frac{\cos^2\theta_1 \cos^2\theta_2}{\sqrt{\lambda L_0}} \frac{\int_{-\infty}^{\infty} \psi_1(x_1)\exp(-\mathrm{i}kx_1\sin\theta_1)\mathrm{d}x_1 \int_{-\infty}^{\infty} \psi_2(x_2)\exp(-\mathrm{i}kx_2\sin\theta_2)\mathrm{d}x_2}{\int_{-\infty}^{\infty} \psi_2^2(x_2)\mathrm{d}x_2}. \tag{3.3.22}$$

如果对式(3.3.21)作旁轴近似: $\cos\theta_1 \to 1, \cos\theta_2 \to 1$,则

$$\eta_{1\to 2} = \frac{a_1 a_2}{\lambda L_0} \frac{S(\theta_1) S(\theta_2)}{\int_{-\infty}^{\infty} \psi_1^2(x_1)\mathrm{d}x_1 \int_{-\infty}^{\infty} \psi_2^2(x_2)\mathrm{d}x_2}, \tag{3.3.23}$$

对应最大振幅传递系数为

$$T_{1\to 2} = \frac{1}{\sqrt{\lambda L_0}} \frac{\int_{-\infty}^{\infty} \psi_1(x_1)\mathrm{d}x_1 \int_{-\infty}^{\infty} \psi_2(x_2)\mathrm{d}x_2}{\int_{-\infty}^{\infty} \psi_2^2(x_2)\mathrm{d}x_2}. \tag{3.3.24}$$

式(3.3.21)考虑了两个波导角度的影响,从理论上说,比式(3.3.23)更为严格地描述了两个法线呈任意夹角的波导间发射和接收能量的耦合效率.

§3.4　SOI 光波导

SOI 指的是绝缘材料上的硅,是一种如图 3.16 所示的三明治结构的硅材料[25~27].基于 SOI 材料的 CMOS 电路具有集成密度高、速度快、低压、耐高温、抗辐照等优点,其性能明显优于常规的硅集成电路. SOI 将成为实现低压、低功耗的国际微电子发展主流技术,具有极其广阔的应用前景.

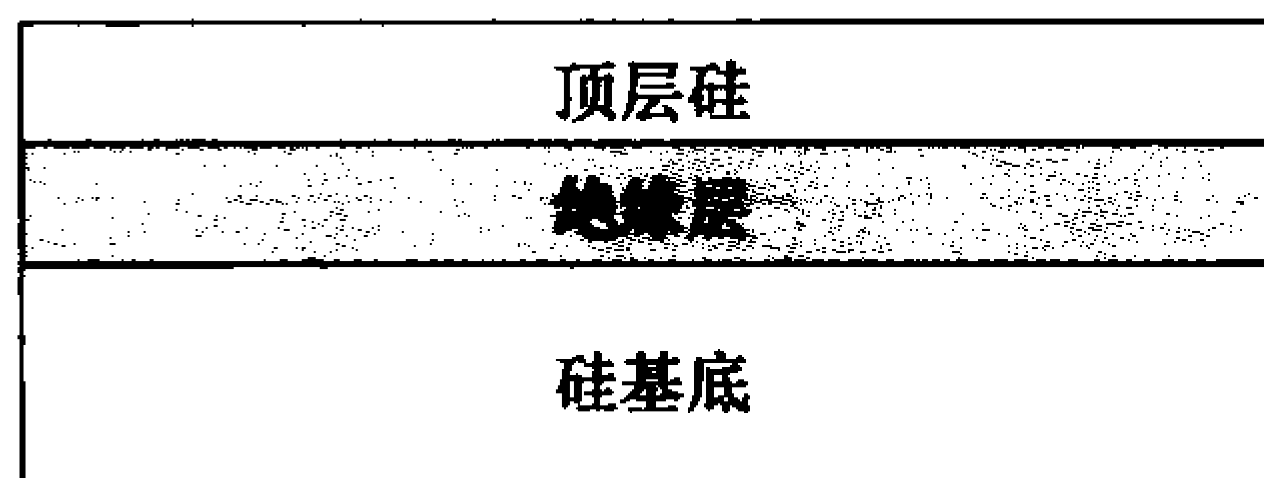


图 3.16　SOI 材料的结构示意图

此外在集成光学领域,SOI 材料也展示了独特的性能优势:由于 Si 和 SiO_2 之间存在很大的材料折射率差别,SOI 材料结构本身就是强限制平板光波导结构,容易获得小尺寸、结构紧凑的 SOI 光波导器件;硅在 1.3～1.6 μm 的光通信波长范围内是“透明的”,光吸收损耗小;而且 SOI 光波导器件工艺与成熟的硅基

CMOS工艺完全兼容，易实现规模化生产. 因此，20世纪80年代开始，将SOI从微电子学领域拓展到光电子学领域，发展集电子学和光子学功能于一体的SOI光波导器件与回路，已经成为硅基光电子集成技术的重要趋势之一.

3.4.1 SOI材料制备工艺与SOI光波导

SOI材料制备工艺的不断完善是SOI光波导器件飞速发展的物质基础. 目前成熟且产业化的SOI材料制备技术主要包括硅片键合和背面腐蚀法(bonding and etch-back SOI, BESOI)、智能剥离技术(smart-cut)和注氧隔离法(separation by implantation of oxygen, SIMOX)，而后两者是最为成功的商业化供片技术[28~31]. BESOI技术是将两片被氧化的硅片键合在一起，然后将其中一片器件硅片(seed wafer)抛光或腐蚀到合适的厚度，另一片支撑硅片(handle wafer)作为机械基底. BESOI材料可以获得初始硅片质量的表层硅和热氧化的SiO_2埋层，但是受减薄技术的限制，表层硅的厚度均匀性难以控制，而且一般只局限于制造微米量级以上的厚膜SOI片. smart-cut技术主要包括氢离子注入、键合、两步热处理、化学机械抛光四个步骤，能够精确控制表层硅厚度，同时提供质量同体硅相同的表层硅，并且能够随意调节SiO_2埋层厚度，但是其工艺复杂，影响最终产品质量的因素多. 而SIMOX技术采用氧离子注入和高温退火工艺在硅片中形成具备理想绝缘埋层的SOI结构，制备过程简单可控，表层硅的厚度均匀性好，工艺条件成熟稳定，成品率高. 虽然SIMOX技术只能制备表层硅厚度为亚微米量级的薄膜SOI材料. 但是通过常规的硅外延技术就可以解决这个问题.

表3.1是SOI材料的主要产业化技术BESOI, SIMOX和smart-cut之间在制备工艺、材料性能与产业化限制因素等方面的比较，不难发现：BESOI技术由于工艺复杂、产品成本高、质量难以保证，将逐渐被淘汰；smart-cut技术和SIMOX技术可以大批量生产出成本相对较低、高质量的SOI圆片. SIMOX技术在表层硅厚度均匀性和制备工艺方面具有一定优势，而且其表层硅与氧化埋

表3.1 SOI材料制备技术的比较

生产技术	SIMOX	smart-cut	BESOI
工艺程序	少	复杂	复杂
表层硅厚度	<400 nm	<2.5 μm	较厚
表层硅厚度均匀性	很好	好	较难控制
埋层厚度	<400 nm	易控制	易控制
界面平整度	易控制	易控制	不易控制
产业化限制	限于注入机	限于专利	产品成本高

层的质量也非常优异，因此 SIMOX SOI 材料是制作 SOI 光波导器件的理想材质.

由于 Si 与 SiO_2 的材料折射率差别大，如果 SOI 平板光波导结构要实现单模传输，则要求波导层的厚度必须小于 0.2 μm，这样在与单模光纤耦合时将引入很大的耦合损耗. 1991 年，Soref 等[15]提出了大截面脊形光波导单模传输理论，通过采用脊形光波导结构，可以使单模 SOI 光波导的截面尺寸扩展到与单模光纤匹配. 图 3.17 是大截面单模 SOI 脊形光波导的剖面图. 随着材料制备工艺和光波导器件加工工艺的不断改善，SOI 单模光波导的损耗不断降低，其损耗可以达到小于 0.1 dB/cm，与 SiO_2 光波导相当. 低损耗 SOI 光波导的研制为 SOI 光波导器件的研制发展奠定了器件结构基础.

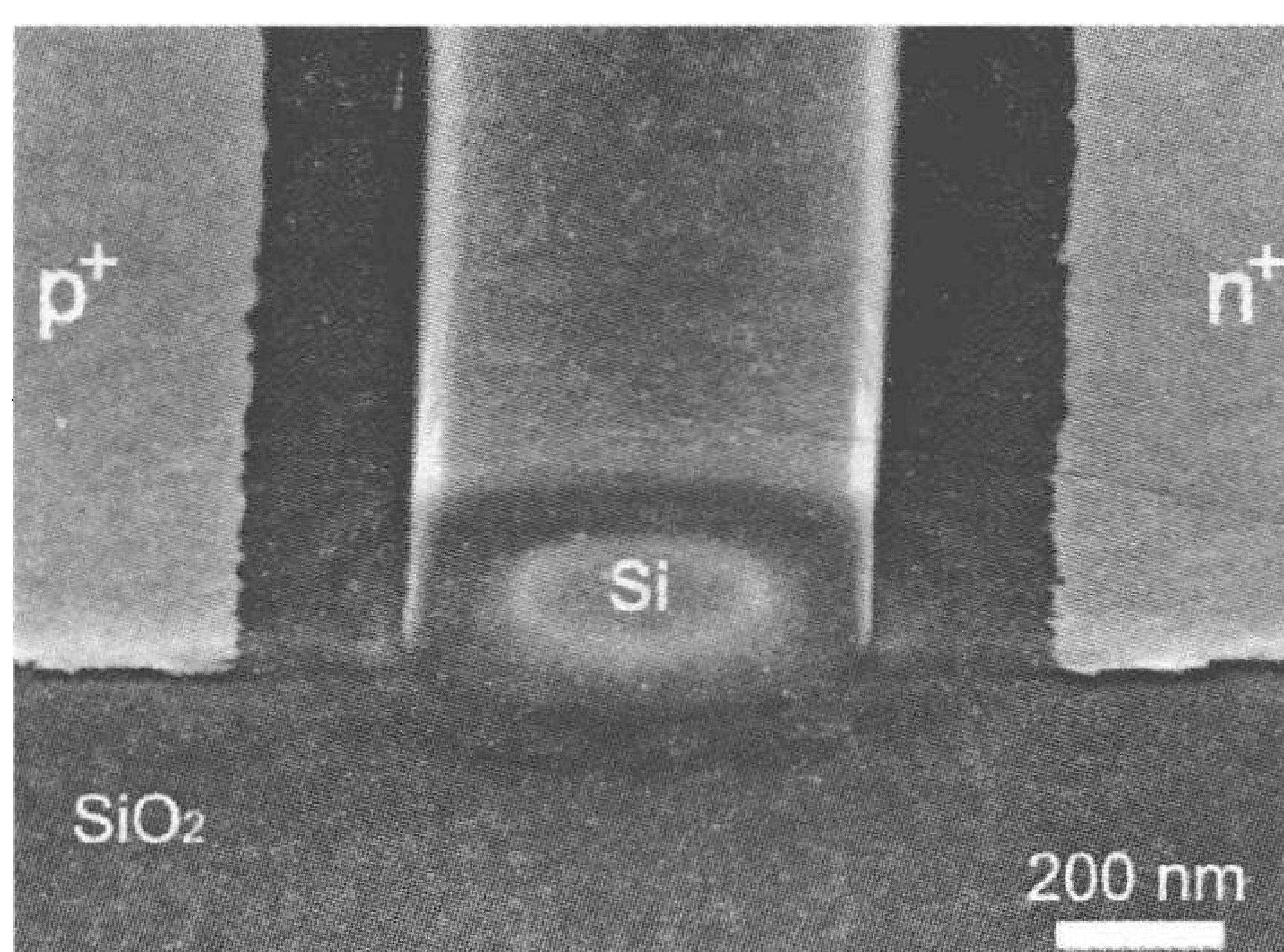


图 3.17　大截面的单模 SOI 脊形光波导的剖面图

弯曲波导在集成光学中是非常重要的. 为了减少器件的尺寸. 希望弯曲波导的弯曲半径尽量小. 波导弯曲时，在弯曲区域传输的导模模式有效折射率将降低，当模式有效折射率小于限制层折射率时光场将向基底泄漏. 这一泄漏损耗机制是限制减少波导弯曲半径的一种重要因素. 由于波导层 Si 和限制层 SiO_2 具有很大的折射率差别，因此这种泄漏损耗在 SOI 光波导中是可以忽略的，容易实现锐弯的光波导结构. 此时限制减小 SOI 弯曲波导的弯曲半径的主要因素是波导侧向辐射损耗. 采用锥形结构(弯曲部分波导的宽度较小)，可以减小弯曲波导的最小弯曲半径.

3.4.2　SOI 光波导器件

低损耗的 SOI 光波导与波导弯曲的研制为 SOI 光波导器件的发展打下了坚实的基础. 在过去的十年里，人们成功地研制了种类繁多的 SOI 光波导无源器件、有源器件和集成回路，并且其中相当一部分已经成功地转化为产品推向市场.

光波导耦合器是集成光学中的一种重要的单元器件. Y 分支波导耦合器以

及输出均衡、插入损耗 1.9 dB 的定向耦合器等传统耦合器都已经成功地制作在 SOI 圆片上，并且利用 SOI 定向耦合器设计、制作了非对称性的马赫-曾德尔(Mach-Zehnder, M-Z)及其级联形式的滤波器. Trinh 等基于 SOI 脊形光波导结构获得了 5×9(输入×输出)的低损耗的星形耦合器. SOI 多模干涉耦合器及其锥形多模干涉耦合器也由 Trinh 等刚研制成功，具有结构紧凑、制作容差大、输出功率均衡的特点.

AWG 是 WDM 通信系统的关键器件. 目前 AWG 的研究主要集中在 SiO_2 和 InP 材料上，然而 SOI 也是一种制作 AWG 的好材料. Jalali 等[12]已经研制了制作在 4 英寸 SOI 圆片上的 AWG，其自由光谱范围(free spectral range, FSR)为 7.6 nm，通道间隔为 1.9 nm，相邻通道间的串扰约为 22 dB，这个器件小于 6 dB. 另外，他们还制作了通道间隔为 2 nm 的 8 通道 AWG.

波导调制器与光开关等是 SOI 基光电子集成技术的重要器件，目前等离子色散效应和热光效应是研制波导调制器与光开关的主要机制. 由于 Si 具有较大的热光系数($\sim10^{-4}$/K)，而且硅的导热性非常好，因此 SOI 热光器件的响应速度比其他材料(如 SiO_2)热光器件的响应快. 实验表明，优化后的 SOI 热光开关的开关时间可以达到微秒量级. 硅是反演对称性的晶体，因此不具备线性电光效应(泡克耳斯效应)，而其非线性(二次)克尔(Kerr)效应也非常弱. 但其自由载流子效应(等离子色散效应)相当显著. 这样靠注入电流可以改变波导中的自由载流子浓度，自由载流子等离子体通过吸收系数对折射率的负作用控制波导折射率的变化，足以实现电光调制. 在 20 世纪 90 年代，SOI 波导电光调制器件取得了长足的进展，如采用全内反射结构的光开关，消光比为 −18.6 dB，插入损耗 4.8 dB；M-Z 干涉型调制器，调制深度为 98%，调制器的响应时间小于 0.2 μs，插入损耗为 4.81 dB. 1995 年，利用 Si 的等离子色散效应在 SOI 脊形光波导结构上实现了电注入 SOI 光波导开关；2004 年，Intel 公司采用 MOS 电容器结构成功解决了硅基电光调制器的响应速度落后于其他波导材料调制器件(如 $LiNbO_3$ 等)的问题，获得了调制频率达到 1 GHz 的 SOI 基集成波导调制器.

光探测器是光集成技术中的重要器件. 尽管硅在通信窗口是透明的，但是采用 GeSi 异质结材料可以实现通信窗口的硅基探测器. GeSi 探测器的结构之一是波导型探测器，采用 SOI 波导结构可以防止能量向 Si 基底中泄漏而导致探测器的响应度比较小. 另一种结构是共振腔增强(resonant cavity enhanced, RCE)型，包括垂直 RCE 和水平 RCE. Cheng 等采用 SIMOX SOI 作底反射器制作了响应峰值波长为 1.285 μm 的垂直 RCE 探测器，反偏压为 5 V 时峰值响应度为 10.2 mA/W，反偏压为 16 V 时在 1.3 μm 测量的外量子效率为 3.5%. 水平 REC 探测器是通过波导上布拉格光栅出现的光子带隙(PBG)效应，利用 PBG 效应的周期结构可以实现高品质因素和大的自由光谱范围的共振腔. SOI 材料

完全符合在三维波导结构中要实现 PBG 效应的折射率调制要求，基于 SOI 的 PBG 结构也已经研制出来.

SOI 光波导器件的发展目的是实现低成本的 SOI 基单片集成光电子回路. 图 3.18 是 B. Jalali 研究小组提出的 SOI 集成光电芯片. 目前进行 SOI 光波导器件研究的机构主要有：西安交通大学、中国科学院半导体研究所、武汉邮电科学院、德国柏林工业大学、美国洛杉矶加州大学等. 英国 Bookham 公司已经可以大批量供应高性能的 SOI 光波导器件，如电调可调式光衰减器，波分复用与解复用器、光发射机、光接收机等等.

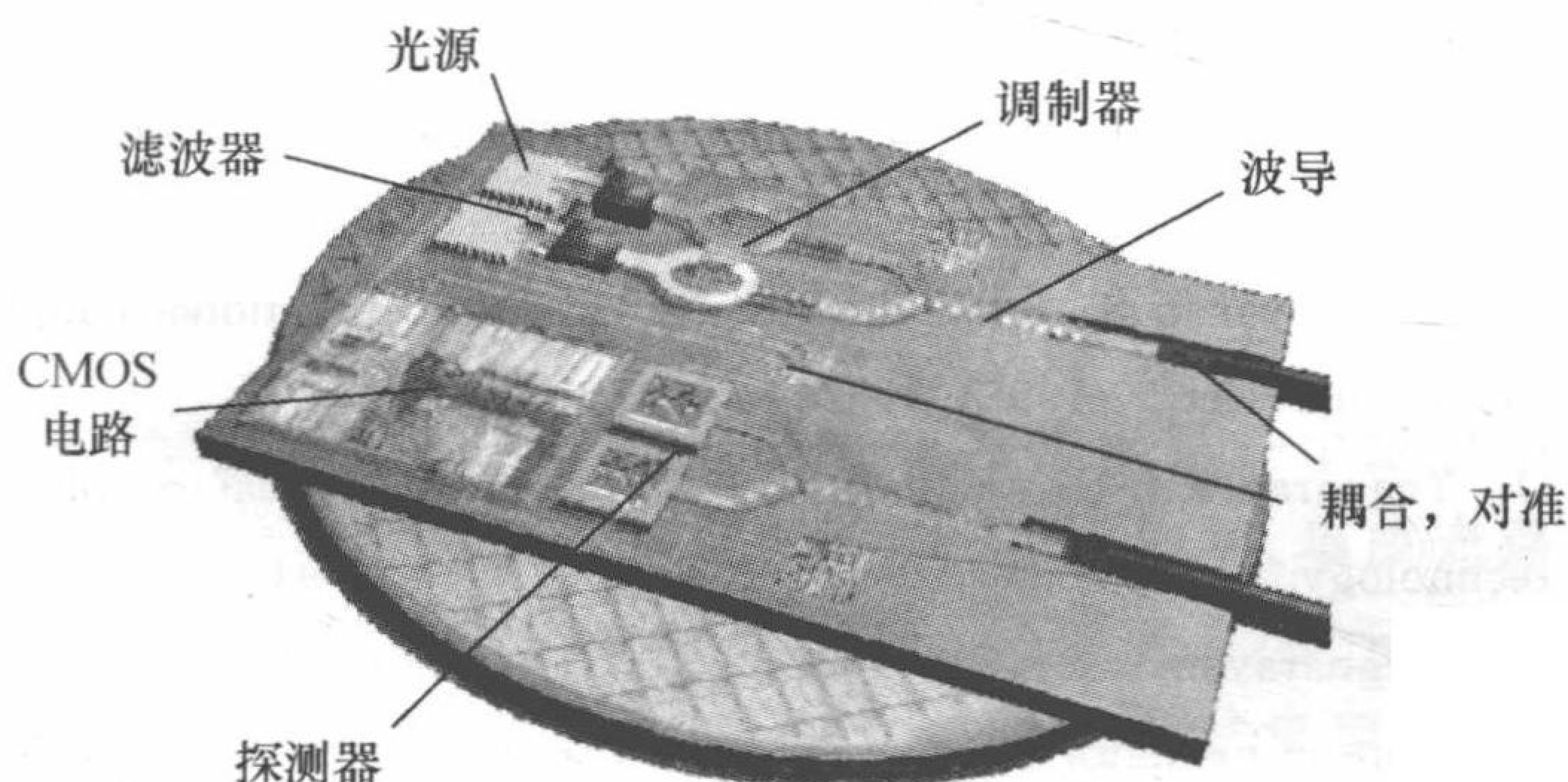


图 3.18　SOI 光电集成芯片

参考文献

[1] 毕德显. 电磁场理论. 北京：电子工业出版社，1985.

[2] 范崇澄，彭吉虎. 导波光学. 北京：北京理工大学出版社，1988.

[3] 玻恩 M，沃耳夫 E. 光学原理. 杨葭荪，等译. 北京：科学出版社，1978.

[4] 陈根祥. 光波技术基础. 北京：中国铁道出版社，2000.

[5] 叶培大. 光波导技术基本理论. 北京：人民邮电出版社，1984.

[6] 胡鸿璋. 应用光学原理. 北京：机械工业出版社，1993.

[7] 方俊鑫. 光波导技术物理基础. 上海：上海交通大学出版社，1987.

[8] Soref R. Applications of silicon-based optoelectronics. MRS Bull.，1998，23：20.

[9] Peters L. SOI takes over where silicon leaves off. Semiconduct or International，1993，16：48.

[10] Soref R A，Lorenzo J P. All-silicon active and passive guided-wave components for $\lambda=$ 1.3 and 1.6 μm. IEEE J. Quant. Electr.，1986，221(6)：873.

[11] Soref R A. Silicon-based optoelectronics. Proc. IEEE，1993，81(12)：1687.

[12] Jalali B，Trinh P D，Yegnanarayanan S，et al. Guided-wave optics in silicon-on-insulator technology. IEE Proc. Optoelectron.，1996，143(5)：307.

[13] Jalali B，Yegnanarayanan S，Yoon T，et al. Advances in silicon-on-insulator optoelectronics.

J. Sel. Topic. Quan., 1998, 4(6): 938.

[14] Day I E, Roberts S W, Knights A. Single-chip variable optical attenuator and multiplexer subsystem integration. Proc. Opt. Fiber Communication Conference and Exhibit, 2002.

[15] Soref R A, Schmidtchen J, Petermann K. Large single-mode rib waveguides in GeSi-Si and Si-on-SiO_2. IEEE J. Quant. Electr., 1991, 27(8): 1971.

[16] Pavesi L. Optical Interconnects: The Silicon Approach. Berlin: Springer, 1998.

[17] Fischer U, Zinke T, Kropp J R, et al. 0.1 dBcm waveguide losses in single-mode SOI rib waveguides, IEEE Photo. Tech. Lett., 1996, 8(5): 647.

[18] Spiekman L H, Oie Y S, Metaal E G, et al. Ultrasmall waveguide bends: the comer mirrors of the future7. IEE Proc. Optoelectron., 1994, 142: 61.

[19] Rickman A G, Reed G T. Silicon-on-insulator optical rib waveguides: loss, mode characteristics, bertts and Y-junctions. IEE Proc. Optoelectron., 1994,141(6): 391.

[20] Trinh P D, Yegnarayanan S, Jalali B. Integrated optical directional couplers in silicon-on-insulator. Electr. Lett., 1995, 31(24): 2097.

[21] Trinh P D, Yegnarayanan S, Jalali B. 5×9 integrated optical coupler in silicon-on-insulator technology. IEEE Photo. Tech. Lett., 1996, 8: 794.

[22] Trinh P D, Yegnarayanan S, Coppinger F, et al. Compact multi-mode interference couplers in silicon-on-insulator technology. Proc. Conf. Lasers and Electro-optics, 1997.

[23] Wei H Z, Yu J Z, Liu Z L, et al. Fabrication of 4×4 tapered multimode interference coupler with large cross section. IEEE Photo. Tech. Lett., 2001, 13(5): 466.

[24] Treyz G V, May P G, Halbout J M. Silicon optical modulators at 1.3 μm based on free-carrier absorption. IEEE Electr. Dev. Lett., 1991, 12(6): 276.

[25] Liu Y L, Liu K L, et al. Novel silicon waveguide switch based on total internal reflection. Appl. Phys. Lett., 1994, 64(16): 2079.

[26] Liu Y L, Liu K L, et al. Silicon 1×2 digital optical switch using plasma dispersion Electr. Lett., 1994, 30(2): 130.

[27] Zhao C Z, Li G Z, Liu E K, et al. Silicon on insulator Mach-Zehnder waveguide interferometers operating at 1.3 μm. Appl. Phys. Lett., 1995, 67(17): 2448.

[28] Liu A, Jones R, Liao L, et al. A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor. Nature, 2004, 427: 615.

[29] Li C, Yang Q Q, Wang H J, et al. $Si_{1-x}G_x$/Si resonant-cavity-enhanced photodectors with a silicon-on-oxide reflector operating near 1.3 μm. Appl. Phys. Lett., 2000, 77(2): 157.

[30] Yablonovith E, Gmitter T J. Photonic band structure: the face-centered-cubic case. Phys. Rev. Lett., 1989, 63(8): 1950.

[31] Foresi J S, Villeneuve P R, Ferreq J, et al. Photonic-bandgap microcavities in optical waveguides. Nature, 1997, 390: 143.

第四章　硅基无源器件

光无源器件是一种不需要借助外部光或电的能量，由自身就能够完成某种光学功能的光学元器件．本章主要介绍一些重要的硅基光子学无源器件，其中包括光栅器件、光子晶体平板波导器件、多模干涉耦合器、阵列波导光栅波分复用/解复用器以及微环谐振腔．

§4.1　光 栅 器 件

自从 1786 年美国天文学家 Rittenhouse 制作出第一个细金属丝光栅以来，至今已有 200 多年了．时至今日，光栅已在天文学、传感、集成光路、光通信和信息处理等领域发挥着重要的作用．就硅基集成光路来说，由于光栅能够灵活实现相位匹配、耦合、光束整形、波长变换[1]等功能且具有结构简单、易于集成、制作工艺简单等优点，它已被广泛地应用于集成光路的有源和无源器件中．

广义上来说，所谓光栅器件是指具有周期性的光学性能（如折射率、透射率等）或空间结构的衍射元器件．本节将重点讲述折射率在一维方向上发生周期性变化的光栅器件．

4.1.1　布拉格条件

布拉格条件是光栅器件中最基本的关系式，它描述的是入射光波矢和各衍射级波矢之间的关系．

在如图 4.1 所示的结构中，设光栅周期为 T，入射和透射介质折射率分别为 n_1，n_2．定义方向沿着 z 向的光栅矢量 $\boldsymbol{K}$，其大小为

$$K = \frac{2\pi}{\lambda T}, \tag{4.1.1}$$

其中 λ 为入射光波长．入射光波矢 $\boldsymbol{K}_{\text{in}}$ 在 z 方向上的分量大小为

$$K_{\text{in},z} = |\boldsymbol{K}_{\text{in}}|\sin\theta = \frac{2\pi}{\lambda} n_1 \sin\theta. \tag{4.1.2}$$

此时，布拉格条件可表示为

$$K_{m,z} = K_{\text{in},z} + mK, \quad m = 0, \pm 1, \pm 2, \cdots, \tag{4.1.3}$$

这里 m 是衍射级的阶数．上式表明，m 级衍射波矢的 z 分量 $K_{m,z}$ 等于入射波矢的 z 分量 $K_{\text{in},z}$ 与 m 倍的光栅波矢大小 mK 之和．

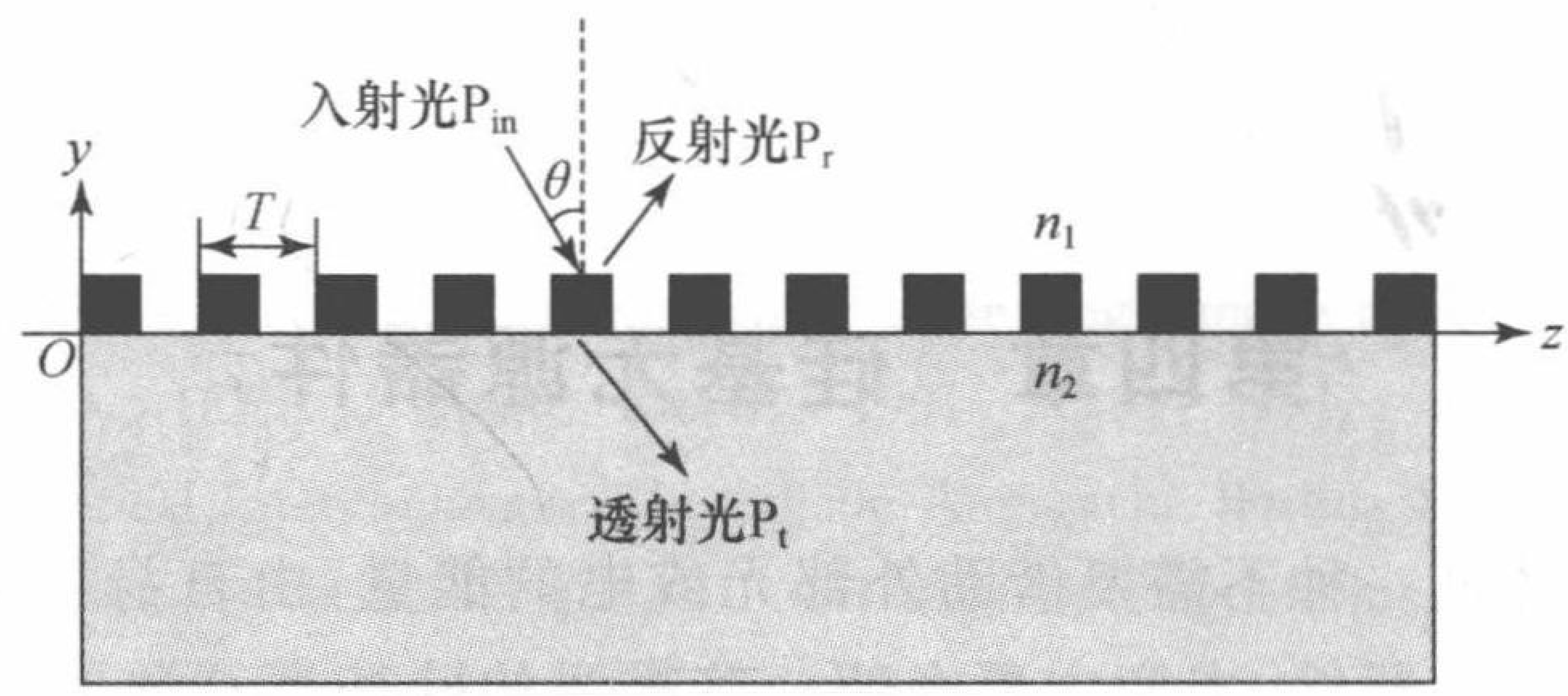


图 4.1 光栅基本结构图

对于上式,还可用矢量关系图来表示,如图 4.2 所示.入射光波矢与光栅矢量的整数倍叠加后所得的波矢为某一衍射波矢.如+1 级透射衍射波矢的 z 分量 $K_{+1,z}=K_{in,z}+K$.

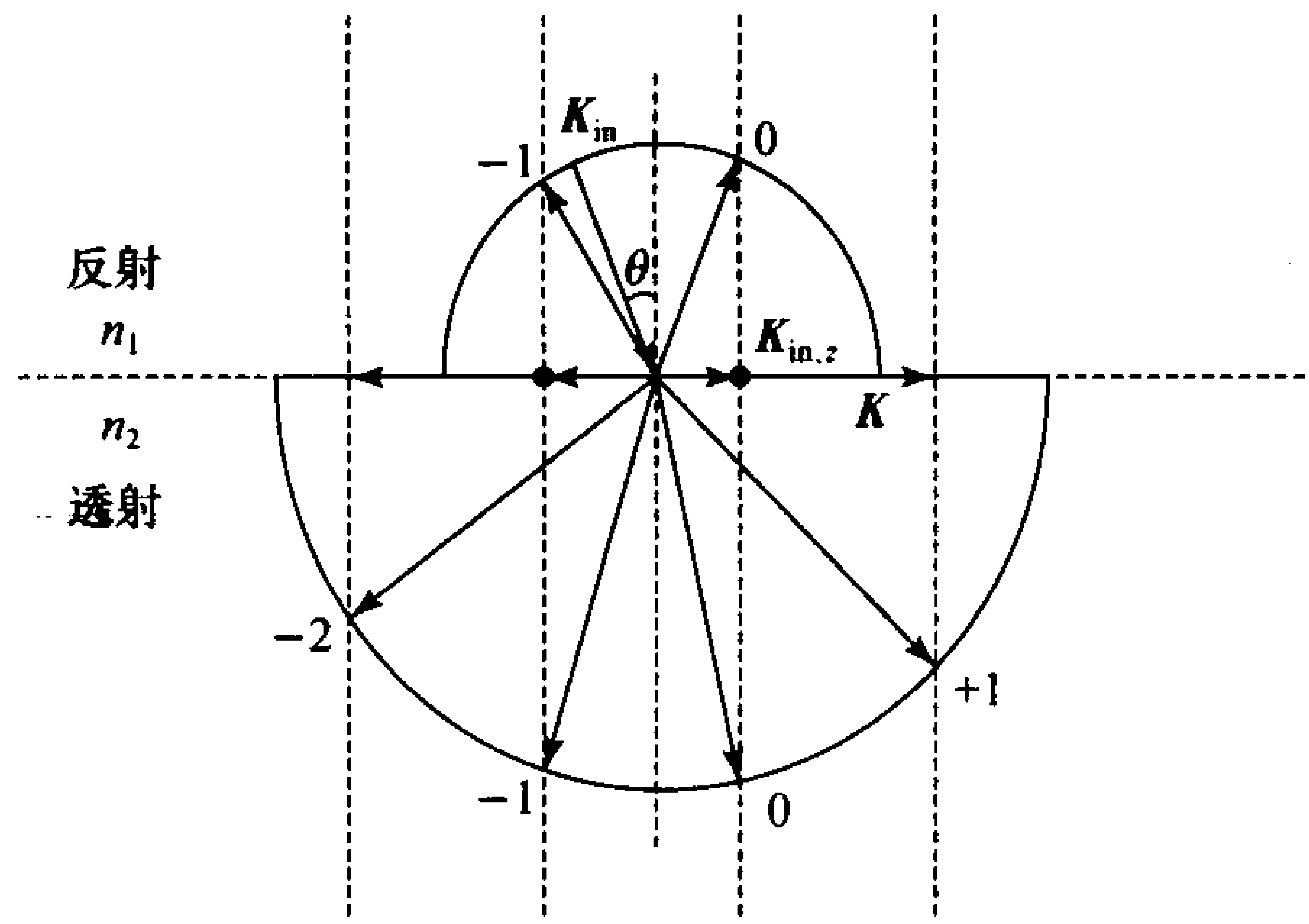


图 4.2 布拉格条件矢量图

4.1.2 光栅器件的应用

1. 光栅耦合器

随着集成光电子器件在光通信系统中的应用,其带来一个问题是,如何将光纤与光电集成芯片波导高效、低成本地对接起来,或者说耦合起来.由于硅基微纳光波导器件中波导的尺寸通常小于 1 μm,而单模光纤的直径为 4~10 μm,光从光纤进入这种小尺寸的波导时,两者之间模斑尺寸以及有效折射率的失配会导致辐射模和背向反射的出现,从而产生很大的插入损耗,所以在集成光学领域,两者之间的耦合问题是一个长期的具有挑战性的课题.

为了降低光纤和芯片波导之间的模式失配和折射率失配损耗,各国研究小组纷纷提出了各种解决方法,如楔形模斑转换器[2~5](spot size converter,

SSC)、反向楔形耦合器[6]、透镜耦合器[7]、棱镜耦合器[8]、光栅耦合器[9~13]等. 然而,楔形耦合器通常制造工艺复杂且难于和其他器件集成,此外输入和输出耦合仅仅发生在集成芯片的边缘,这限制了光输入/输出(input/output, I/O)端口的位置[14]. 透镜和棱镜耦合器难于和其他器件集成且体积较大,不适于现代光电集成的发展方向. 而对于光栅耦合器(参见图 4.3),其耦合可以发生在集成芯片平面内外的任何位置,可以灵活布置 I/O 端口位置,采用这种耦合方式芯片集成度能大大提高,并且可以在圆晶切割成芯片之前进行片上测试. 光栅耦合器主要是通过光栅的衍射作用把光纤中的光有效地耦合进波导,可以采用垂直耦合结构也可以采用平行耦合结构. 理论上光栅耦合器具有很高的耦合效率,是一种非常有前途的耦合方法.

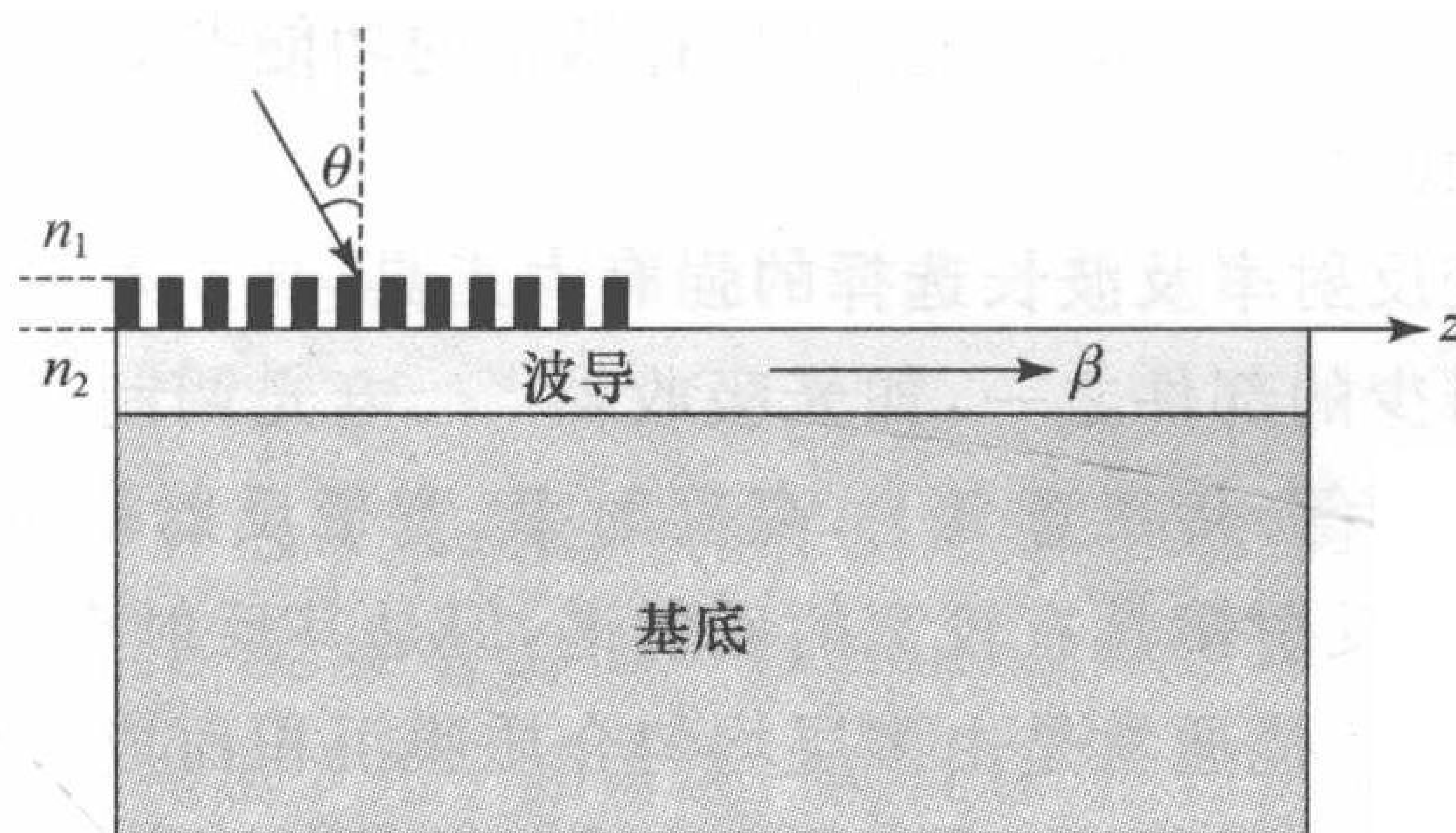


图 4.3　光栅耦合器结构示意图

接下来,讨论布拉格条件在光栅耦合器中的表现形式. 对于图 4.3 所示典型光栅耦合器结构示意图,设波导中导模的传播常数为 β,则此时布拉格条件可表示为

$$\beta = K_{m,z} = K_{\text{in},z} + mK, \quad m = 0, \pm 1, \pm 2, \cdots. \tag{4.1.4}$$

上式表示,如果衍射波的传播常数 $K_{m,z}$ 等于导波传播常数 β 时,那么 m 阶衍射波就可以耦合进波导而成为导模. 相反地,波导中的导模也可以反过来成为 m 阶衍射波而射出.

将式(4.1.1),(4.1.2)代入上式,有

$$\frac{2\pi}{\lambda} n_1 \sin\theta + m \frac{2\pi}{\lambda T} = \beta, \quad m = 0, \pm 1, \pm 2, \cdots. \tag{4.1.5}$$

当光从光纤中输入,此时 n_1 就为光纤折射率,利用光栅就能实现光纤与微纳平面波导的耦合.

这里需要说明的是,布拉格条件的使用是有其限制条件的. 首先,布拉格条件只能说明哪些衍射级可能发生耦合,而不能解决各衍射级能量分布的问题,更不能说明光栅耦合效率问题;其次,布拉格条件只能简单说明光栅耦合时,入射

光波长、入射角、光栅周期和折射率等因素之间的关系，而影响耦合的其他一些重要参数，如光栅深度、占空比、光栅长度等参数都没有涉及.

因此，布拉格条件只能作为设计光栅耦合器的参考，而更为精确的计算需要用到严格耦合波分析(rigorous coupled-wave analysis, RCWA)[15]或时域有限差分(finite-different time-domain, FDTD)[16]等算法.

对于光栅耦合器的研究，比较早的是 M. L. Dakss 和 L. Kuhn 等利用光刻胶制成的光栅耦合器[17]，随后在微纳加工技术的带动和影响下，人们纷纷设计和制作各种不同的光栅耦合器件. 如为了提高光纤与微纳波导的耦合效率，美国阿拉巴马大学的 B. Wang 和 G. Nordin 等提出了倾斜光栅耦合器[14, 18]. 2007 年，国内有课题组报道了一种基于二元闪耀光栅的偏振分束耦合器，它既能将非偏振光分解成两束正交的线偏振光，又具有耦合的功能[19].

2. 光栅反射镜

作为一种可变反射率及波长选择的强有力工具，高反射率、宽带反射镜是许多光学元件必不可少的部件之一，如光吸收器[20]、电光调制器[21]、激光器[22,23]、探测器[24]、耦合器[25]等. 传统意义上，高反射率、宽带反射镜主要是由金属或多层介质膜构成. 金属反射镜具有较宽的反射谱，但是其反射率受限于金属的吸收损耗；多层介质膜反射镜通常是由多层均匀介质膜沉积而成，由于受到介质膜的折射率以及膜厚等因素的影响，其反射率通常不高[26]. 近年来，一种基于介质光栅的泄漏模谐振(leaky mode resonance, LMR)的光栅反射镜[27]受到了世界各国科研小组的关注.

此处的泄漏模谐振是指当入射光与光栅中的泄漏模相位相匹配时，在光栅反射谱中，其 0 级反射率接近 100%的现象[28]. 图 4.4 为光栅未发生和发生泄漏模谐振时的衍射图谱. 对比图 4.4(a)，(b)可知，发生泄漏模谐振时，在衍射图谱中，我们可以观察到 0 级衍射被增强，其反射率将近 100%，而其他高衍射级被压制. 能利用这一特性，光栅可实现高效率的反射滤波. 相对于多层介质膜和光纤光栅滤波器，其易于与光电芯片集成，且制作工艺简单，改变光栅的调制度和形状，它能够灵活地实现窄带和宽带滤波.

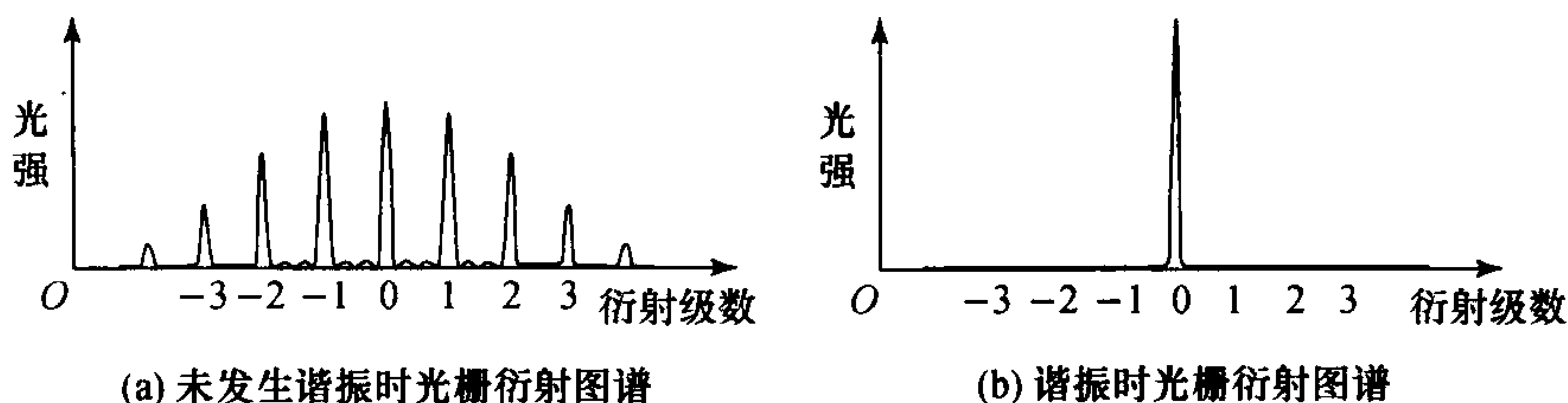


图 4.4　未发生及发生泄漏模谐振时光栅衍射图谱

根据泄漏模谐振原理,2004 年美国伯克利加州大学的 M. C. Y. Huang 和 C. J. Chang-Hasnain 等[26,29,30]利用在硅基底上沉积低折射率包层(cladding)的亚波长光栅,成功设计和制作出了反射响应平坦反射镜,它能在较宽的波段范围内表现出较高的反射率.

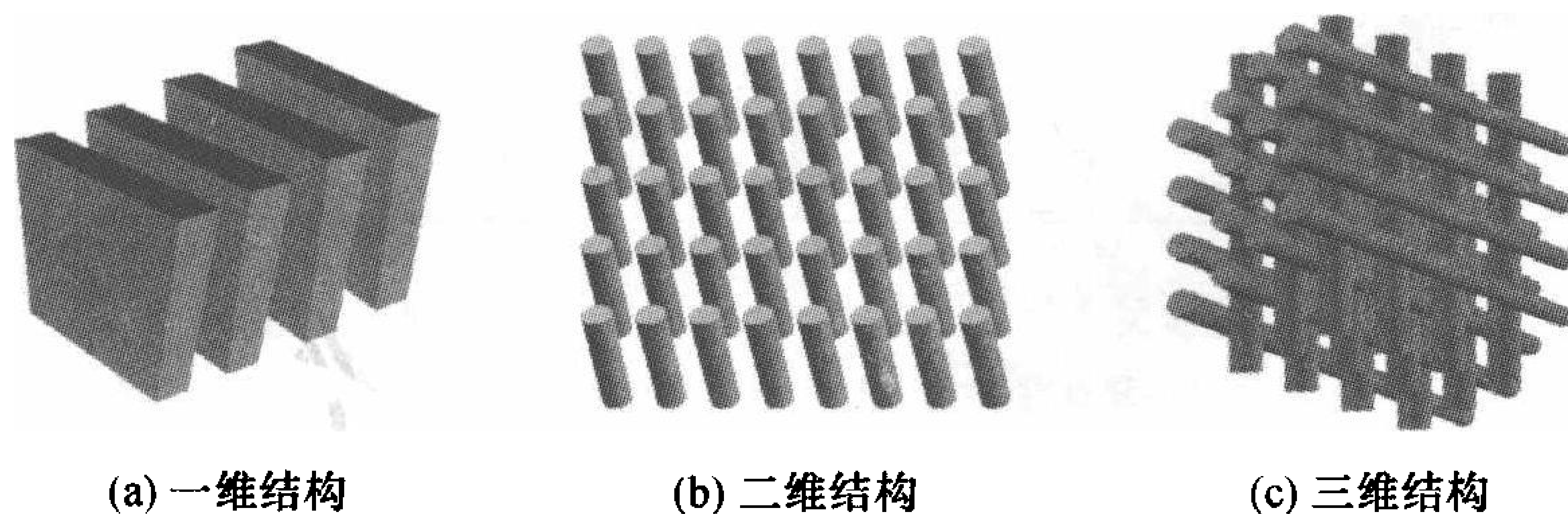
前面介绍的宽带光栅反射镜只能在 TM 偏振状态下实现宽带、高反射率,而对于 TE 偏振光,其性能就大不一样了,这限制了器件的进一步应用. 为了解决这个问题,最近国内有课题组提出了一种偏振无关宽带光栅反射镜[31]. 一束平面单色光垂直照射到此结构上,通过多层光栅结构,TE 和 TM 偏振光在 0 级都被高效的反射,从而达到了偏振无关的目的.

§4.2　光子晶体平板波导

4.2.1　光子晶体基本概念

光子晶体是一种介电常数或磁导率在空间呈周期性分布的人工电磁晶体材料. 1987 年,美国 Bell 实验室的 E. Yablonovitch[32] 和普林斯顿大学的 S. John[33],各自独立地提出了光子晶体这一新概念. 他们发现,与电子在固态晶体的周期性势场下形成的电子能带和带隙类似,光子晶体对光的布拉格散射,也能够在一定频率范围内产生光子带隙. 频率处在光子带隙中的光不能在光子晶体中传播,而几乎完全被反射回去.

光子晶体按空间分布,可以分为一维、二维和三维光子晶体,如图 4.5 所示. 通常,人们习惯把折射率在一维方向上发生周期变化的一维光子晶体称作光栅,如现在广泛应用的多层布拉格光栅结构就是一维光子晶体[34,35];二维光子晶体在实际中通常演变为有一定厚度的光子晶体平板,由于这种结构适合用传统的微电子平面加工工艺来实现,因此成为目前光子晶体器件采用的主要结构[36~38];三维光子晶体具有完全真正带隙,即落在禁带频率范围的光无论从任



(a) 一维结构　(b) 二维结构　(c) 三维结构

图 4.5　光子晶体结构示意图

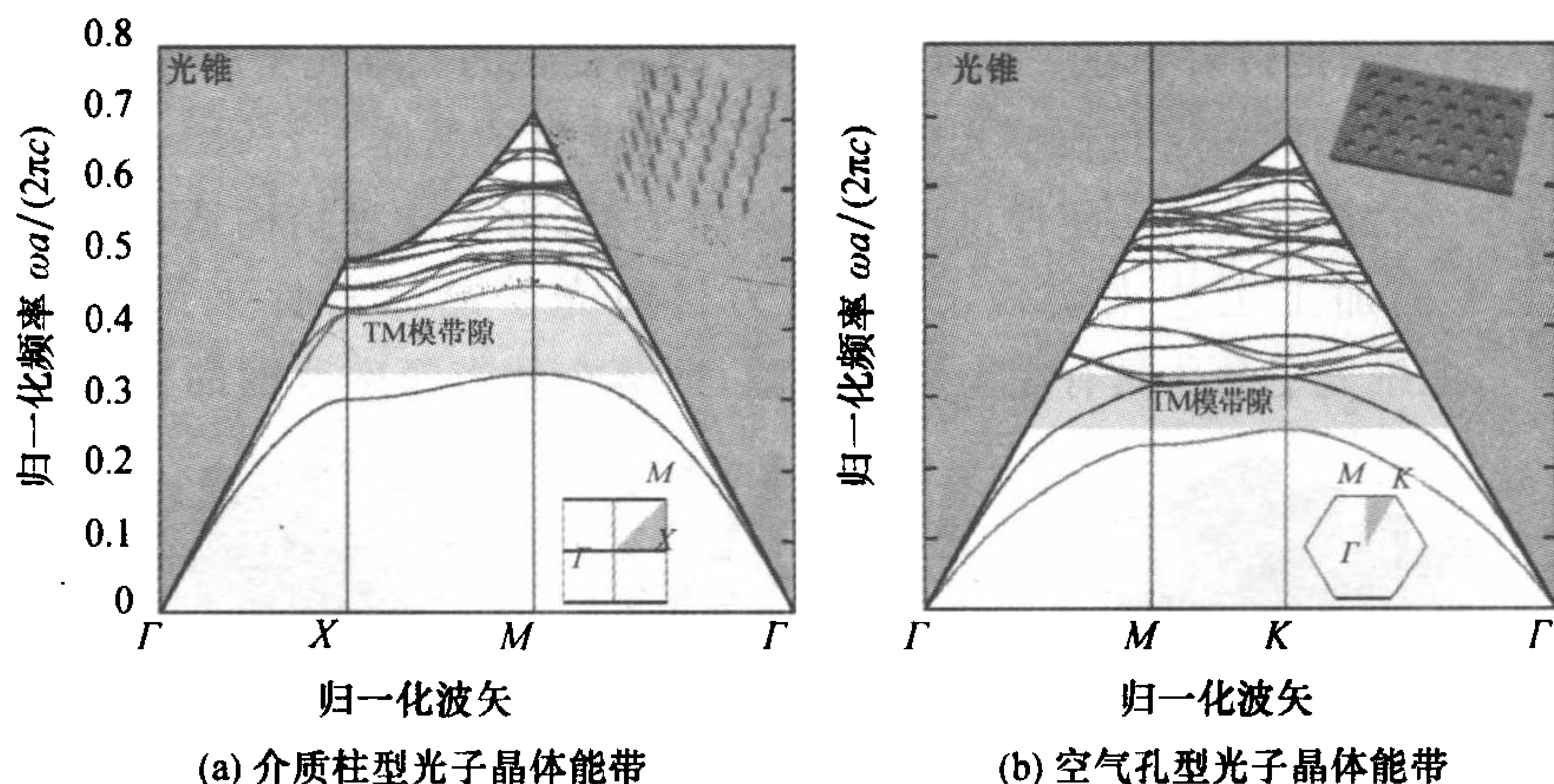
何角度入射都无法在光子晶体内部传播[39~41]. 从理论上说,三维光子晶体的性能最佳,但是目前还没有比较好的制作加工手段.

经过二十多年的发展与完善中,目前光子晶体已广泛应用于微波通信[42,43]、光子芯片[44,45]、太阳能电池[46,47]和生物化学传感[48,49]等技术中.

4.2.2 光子晶体平板

如上文所述,具有一定厚度的二维光子晶体平板能够实现光子晶体的多数特性要求,且可用标准的平面波导加工技术制作,因此成为目前光子晶体器件采用的主要结构.

通常光子晶体平板在周期平面内是利用光子带隙来限制光场的,而在垂直方向上则是利用介质的折射率差来限制光场,利用 SOI 结构的强限制型波导,能够大大减小垂直方向的损耗. 典型的光子晶体平板有两种结构: 介质柱型和空气孔型. 图 4.6(a),(b)的插图分别给出了两种典型的光子晶体平板结构示意图[50]. 从图 4.6 中可以看出,两种结构分别为 TM 场与 TE 场,均存在禁带. 需要特别说明的是,在光子晶体中,TE 与 TM 的电磁场矢量定义与经典光学理论中的电磁场矢量定义是不一样的. 为方便分析研究,习惯上在光子晶体中,认为电场在周期平面内的偏振态均为 TE 的,而磁场在周期平面内的偏振态均为 TM 的[51]. 对于光子晶体平板结构,由于在垂直方向上缺乏平移对称性,所以没有纯粹的 TE 和 TM 模式,但是根据光子晶体平板的水平对称性,可以将模式区分为偶对称和奇对称,这两种模式和二维情况下的 TE 和 TM 模式非常相似,所以偶对称和奇对称模式又分别被称为 TE 和 TM 模式[52].



(a) 介质柱型光子晶体能带　　(b) 空气孔型光子晶体能带

图 4.6 两种典型光子晶体平板能带图

4.2.3　光子晶体平板波导

在完美光子晶体平板中引入一排或几排一维的扰动，即在光子晶体中形成了一维的线缺陷，就形成了光子晶体平板波导[53]. 频率处于光子晶体禁带内的光子，被禁止进入线缺陷周围的光子晶体区. 如果这个线缺陷支持的光模式正好处于光子晶体的禁带频率内，由于光子晶体的光子带隙作用，该频率范围内的光进入缺陷区域后会被光子晶体区反射而只能沿线缺陷方向传播. 如图 4.7 所示，在硅基二维三角晶格光子晶体中移除一排空气孔形成的光子晶体平板线缺陷波导. 由于两旁光子晶体的存在，频率处于禁带内的光只能被限制在中央的波导中传输，而不会扩散到两边. 移除 n 排平行的空气孔，将得到一个 W_n(n 排空气孔宽度)光子晶体线缺陷波导. 一般来说，缺陷的宽度越宽，光子晶体缺陷波导支持的模式越多，但多个模式之间很容易因模式间的相互耦合而产生色散，导致光信号传送失真. 因此，光子晶体波导通常都是工作在单模状态.

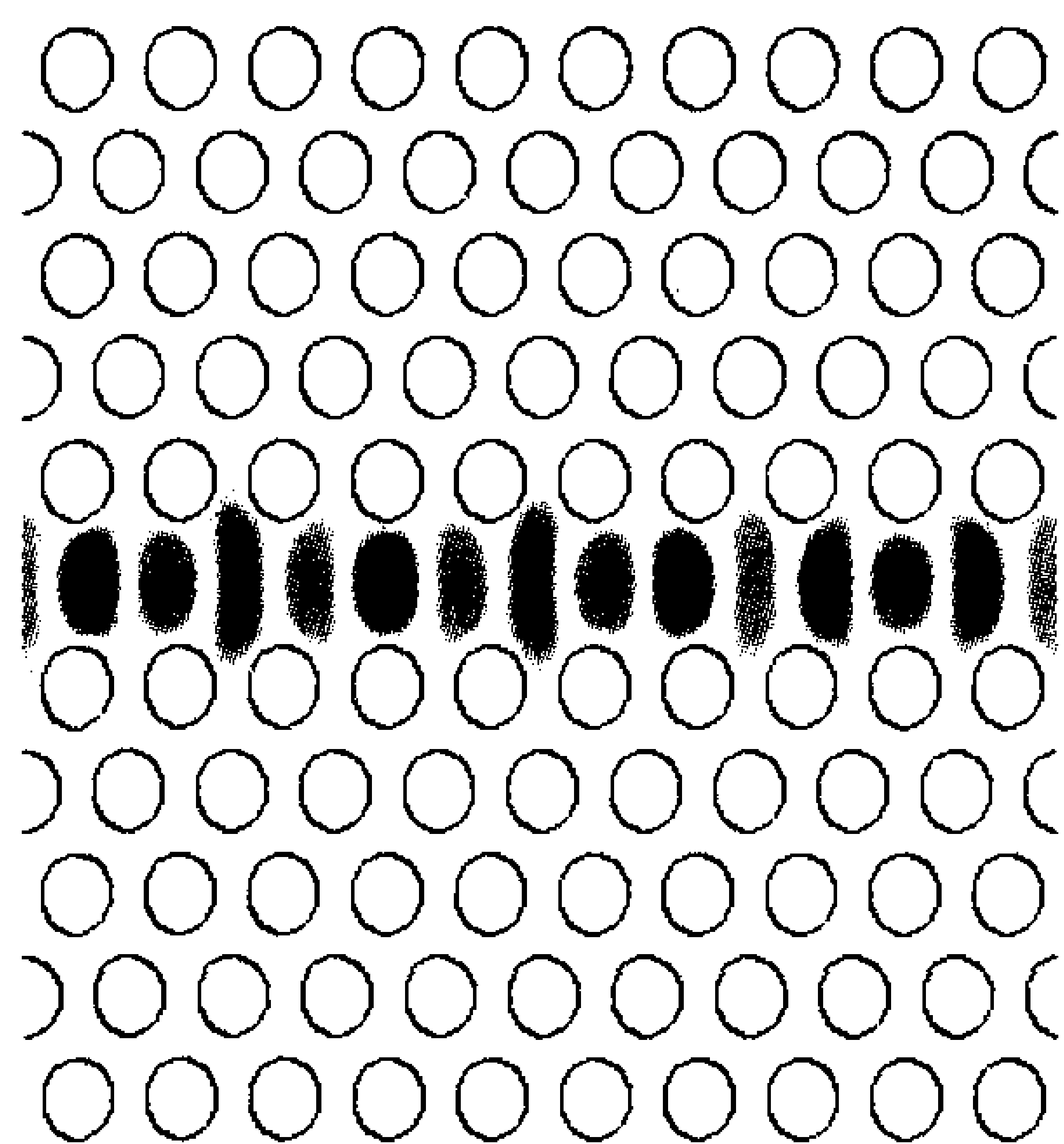


图 4.7　光波在硅基光子晶体平板单线缺陷波导中传播

图 4.8 是图 4.7 所示结构的能带结构图. 图中横坐标表示的是归一化轴向传播常数 $\beta=k_x=ka/(2\pi)$，其中 a 是晶格常数. 纵坐标是归一化频率 $\omega a/(2\pi c)=a/\lambda$，$\lambda$ 是光波长. 如图所示，对于光子晶体平板波导，反映垂直光限制的色散曲线，通常被称为光线，光线以上的区域称为光锥. 位于光锥中的模式称为辐射模，其在光子晶体中传播的时候，光会泄漏到包层中去，损耗很大；而在这条线以下的模式，如灰色区域中的光子晶体平板模式，只能在光子晶体中传播，不会泄漏到包层中去；处于禁带中的模式，如 1 阶导模与 2 阶导模，只能被束缚在缺陷中传播，损耗几乎为零.

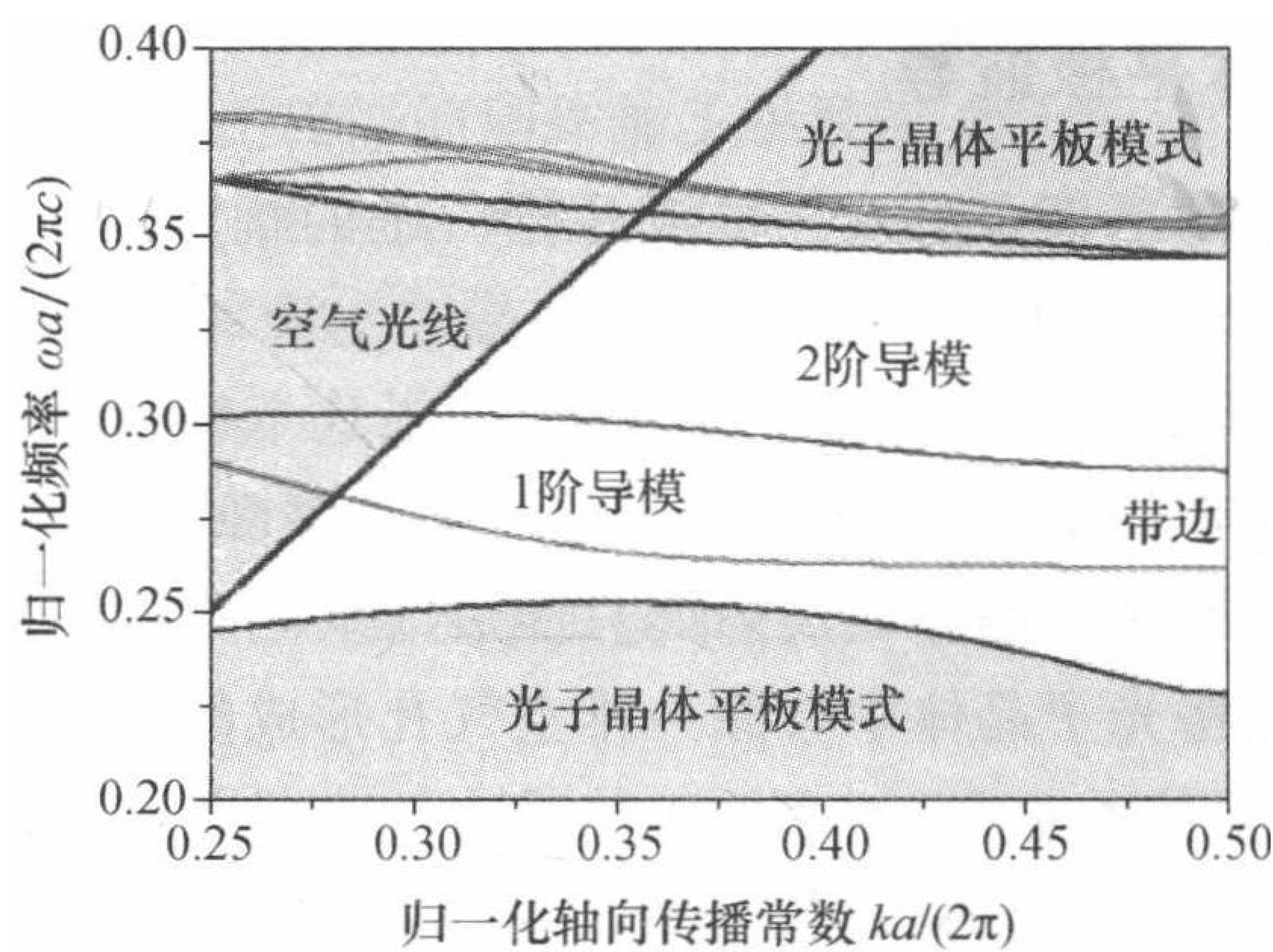


图 4.8 硅基光子晶体平板单线缺陷波导能带结构图

4.2.4 光子晶体波导器件

1. Y 分支功分器

基于传统波导的 Y 分支功分器的分束角受辐射损耗的限制，通常在几度以内，因此为了将能量输出到两个分开的波导中，整个器件的横向尺寸比较大. 而基于光子晶体波导的 Y 分支功分器的分束角可以达到 120°，从而能够大大地减小器件尺寸. 图 4.9 所示为丹麦技术大学等提出的基于 SOI 结构的光子晶体平板波导 Y 分支功分器[54]. 其基本光子晶体波导结构为单排线缺陷光子晶体波导，Y 分支功分器由 120°Y 分支和两个 60°的弯曲波导组成. 实验测得：在 1560～1585 nm 波长范围内，其 3 dB 输出损耗几乎为 0. 整个器件的尺寸为 15 μm×20 μm，且该器件还表现出较好的工艺容差性.

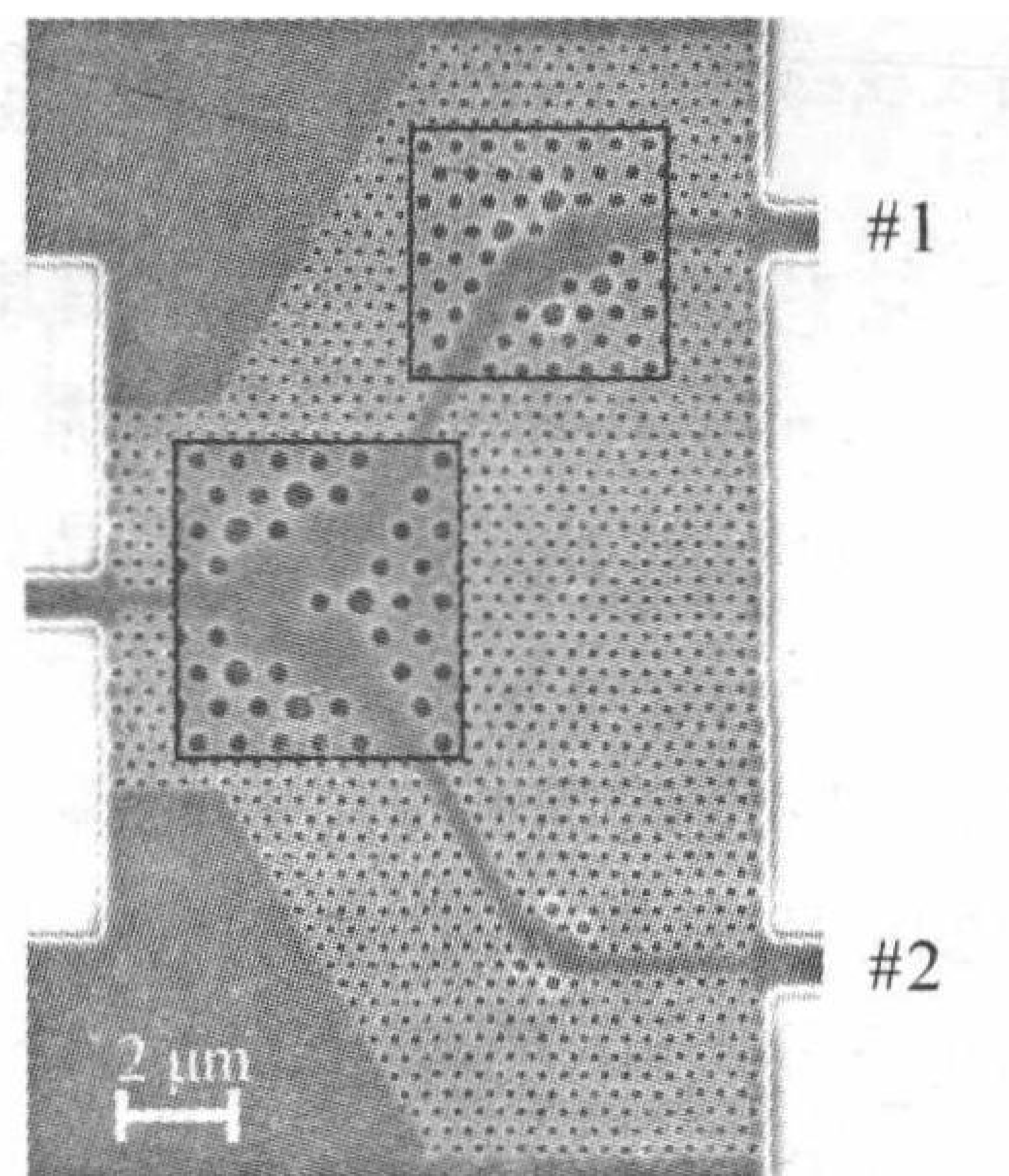


图 4.9 SOI 结构光子晶体平板波导 Y 分支功分器

2. 光子晶体慢光波导

所谓慢光是指光速远远小于光在真空中的传播速度($c\approx3\times10^8$ m/s)的一种现象.慢光在光学延迟线、光缓存和光计算等领域都有潜在的应用.通过对光子晶体结构的优化调节,许多慢光领域的主要问题可以得到解决,如:光速的减慢与频率带宽的增加之间的矛盾,由强色散导致的光信号变形,以及输入、输出端口上的反射损耗等.因此,光子晶体平板波导被认为是微纳结构慢光主要实现方式之一.

首先,来看看慢光产生的机理.对于折射率为 n 的介质,其中的光速可以描述为 c/n.但对于一个光脉冲信号来说,就要用群折射率 n_g 代替介质折射率 n,相应的就存在群速度 v_g,其定义为

$$v_g=\frac{\mathrm{d}\omega}{\mathrm{d}k}=\frac{c}{n+\omega\dfrac{\mathrm{d}n}{\mathrm{d}\omega}}=\frac{c}{n_g},\tag{4.2.1}$$

其中,$n_g=n+\omega\dfrac{\mathrm{d}n}{\mathrm{d}\omega}$为群折射率.$\dfrac{\mathrm{d}n}{\mathrm{d}\omega}$表示材料的折射率与入射光频之间的关系,也称色散关系.由上式可知,要获得慢光,也就是较低的群速度,可以通过增大$\dfrac{\mathrm{d}n}{\mathrm{d}\omega}$来实现,比如引入较大的色散.

2005 年,IBM 公司 Watson 研究中心的 Y. A. Vlasov 等[44]利用集成在硅芯片上的低损耗硅光子晶体波导,在横截面加上有源电极,通过电流的改变来调节介质折射率,从而达到改变群速度的目的,其扫描电镜(scanning electron microscope, SEM)的照片如图 4.10 所示.通过实验,他们得到了速度约为光速的 1/300 的群速度(c/300),使基于光子晶体的慢光向实际应用迈出了关键性的一步.但是这种结构存在着群速度色散大(约为传统光纤的 10^7 倍),在传输过程中会极大的展宽脉冲,造成严重信号失真的问题.

随后,日本横滨国立大学的 Mori 和 Baba[55]设计了一种啁啾结构的光子晶体耦合波导,它能在带隙范围内形成一条拥有拐点的“座椅状”平坦能带.仿真结果表明,此波导在慢光传输模式下,相比传统的线缺陷光子晶体波导具有更小的群速度色散.2007 年,台湾新竹交通大学和日本 NTT 公司 Huang Shih-Chieh 等[56]对此结构进行了实验验证,以及时域和频域的性能考察.通过调整输入脉冲的波长,他们将群速度降为 0.017c.

“座椅状”形式的光子能带结构的好处是,不仅能使得慢光传播过程中的群速度色散获得补偿,而且由于光子能带的慢波区域远离布里渊区边界,其上下都不存在截止频率,从而有效降低了反射损耗.实际中,通常是通过啁啾结构上下移动光子带,这不仅补偿了单一频率的群速度色散,也同时增加了慢波的带宽.

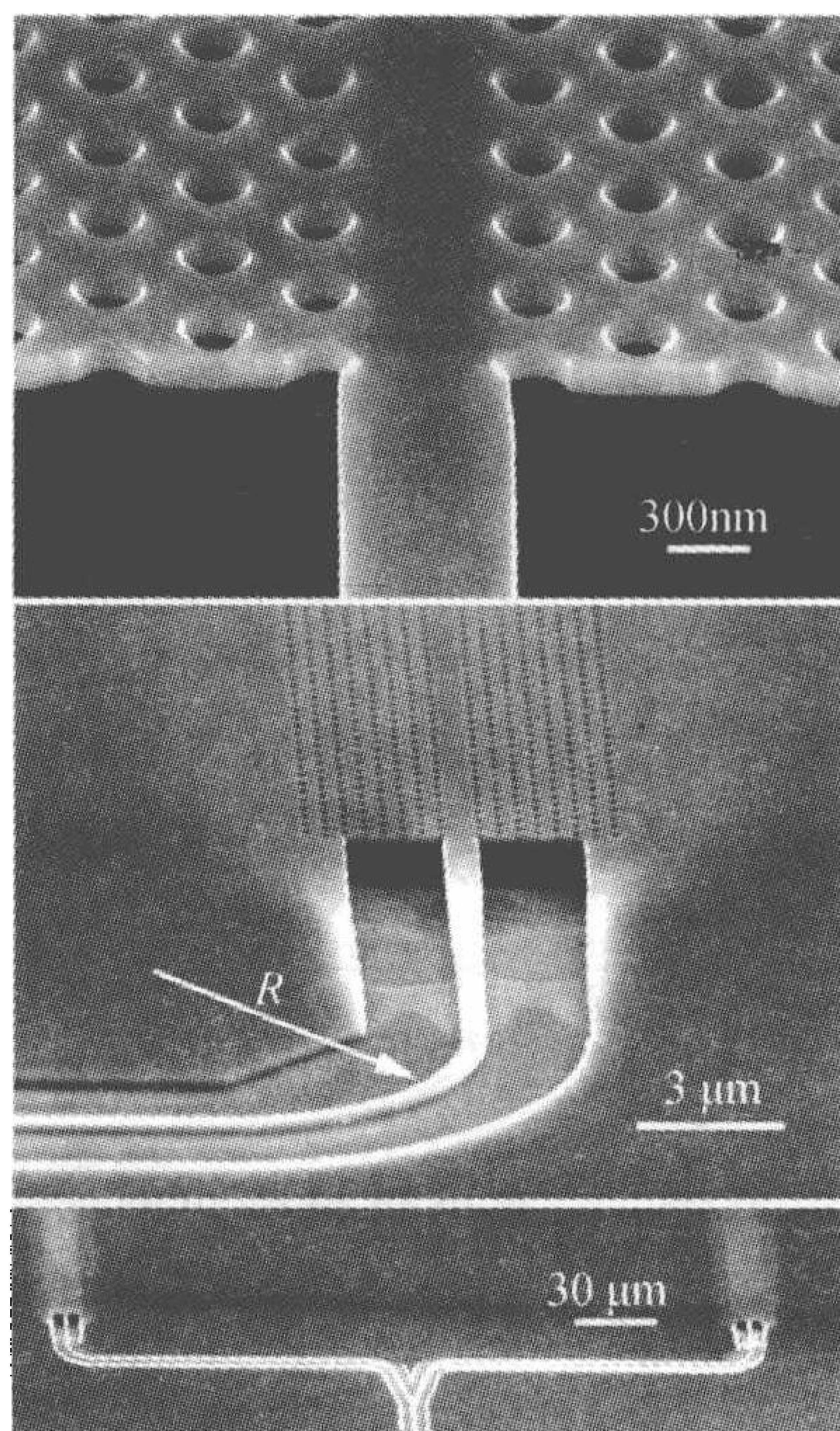


图 4.10 IBM 公司 Watson 研究中心实现的 M-Z 干涉结构图

北京大学课题组[57]通过调整结构的方式，在最接近波导的两排空气孔中增加内介质柱，通过改变内介质柱的直径达到调整色散曲线形状的目的，从而实现了宽带慢光. 随后，基于啁啾刻槽结构的光子晶体耦合波导，该课题组[58]提出了一种光子能带的慢波区域远离布里渊区边界的结构. 当结构优化后，可得归一化频宽 $\Delta\omega/\omega=1.953\,32\%$，即中心波长 1550 nm 处带宽为 30 nm.

虽然光子晶体平板波导是实现慢光的有效方式之一，然而受现有实际工艺条件限制，其制备工艺复杂，制作成本高；而一维周期性介质波导结构简单，易于加工，方便与其他器件集成，因而越来越受到国内外研究者的关注.

2009 年，上海交通大学课题组[59]设计了一种椭圆形空气孔阵列的直波导. 通过参数优化后，得到了一条平坦的色散曲线. 经计算后，得到归一化频宽 $\Delta\omega/\omega=0.0396\%$，平均群折射率 $\bar{n}_g=418$. 同时，该课题组还设计了一种新型的沙漏形空气孔阵列的直波导，这种空气孔由上、下两个对称的半椭圆组成. 经计算后，得到归一化频宽 $\Delta\omega/\omega=0.212\%$，平均群折射率 $\bar{n}_g=87.8$.

综上所述，由于慢光在全光信号处理、光计算、非线性增强以及光学传感等方面具有重要的应用价值，而硅基微纳结构的慢光技术是实现单片集成、发挥慢光重要作用的有效手段之一，因而硅基微纳结构的慢光实现技术将是当前和今后国内外信息技术研究的热点之一.

3. 光子晶体调制器

传统的硅基调制器基于脊波导和 M-Z 干涉仪(MZI)结构，由于脊波导的相位调制效率较低，通常 MZI 调制臂长度需要几毫米．而光子晶体慢光波导具有极低的群速度和强色散，因而用其充当调制臂可以大幅缩短臂长，而调制臂的缩短使得电极长度也得以缩短，从而导致电功耗也显著降低．2004 年，C. Liu 等从理论上设计了臂长仅为 10.26 μm 的光子晶体 MZI 调制器，损耗仅为 1%，但响应速度只达到毫秒量级[60]．随后，美国德州大学奥斯汀分校的 C. Ray 研究组设计并制作了 80 μm 长的光子晶体 MZI 调制器[61]，测得的调制深度达到 92%.

§4.3　多模干涉耦合器

近年来，基于平面波导型的多模干涉(multimode interference，MMI)耦合器件，由于较其他器件具有插入损耗低、频带宽、制作工艺简单和容差性好等优点，已被广泛地应用于光开关[62]、波分复用/解复用器[63]、功率分配和组合器[64]等重要光学器件中．MMI 耦合器根据端口数量，主要有 $1\times N$ 型和 $N\times N$ 型两种．目前对 MMI 耦合器的研究主要集中于研究多模波导的结构，减小器件尺寸，以及提高功分器的性能.

4.3.1　自镜像效应

在多模干涉耦合器的设计过程中，人们通常用自镜像效应(self-imaging effect，SIE)来确定器件的结构参数．对于自镜像效应，可以这样简单地解释为：自镜像效应是多模波导的一个重要特性，它是波导中被激励起来的模式间的相长性干涉的结果．通过这个效应，沿导波的传播方向将周期性地产生输入场的一个或多个像[65]．下面就以图 4.11 所示 $N\times N$ 型 MMI 耦合器为例，对自镜像效应加以说明.

如图 4.11 所示，该器件由输入波导、多模干涉区以及输出波导三部分组成.

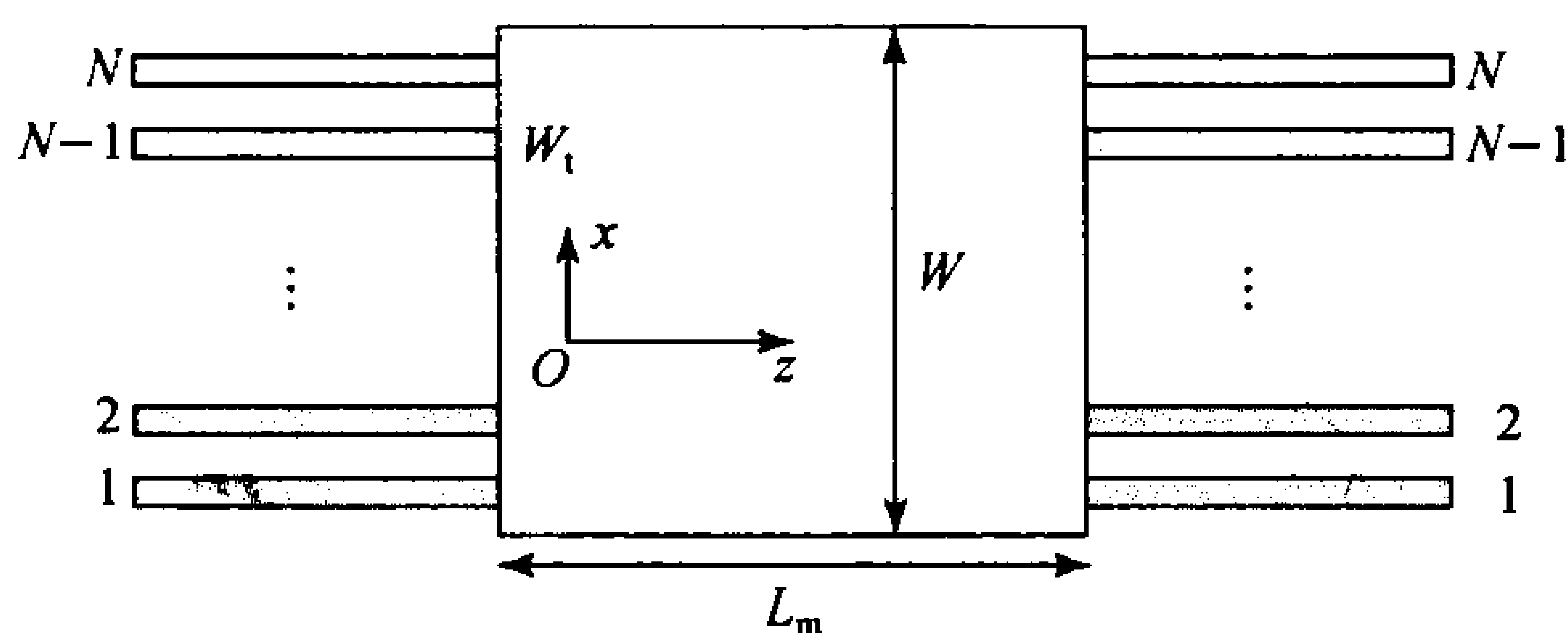


图 4.11　$N\times N$ 型 MMI 耦合器结构示意图

设 MMI 部分的宽度和长度分别为 W 和 L_m，其心层和包层的有效折射率分别为 n_r 和 n_c，输入/输出波导的宽度为 W_t.

在 MMI 部分的输入处($z=0$)光场可以展开为

$$f(x,z=0)=\sum_m c_m\varphi_m(x), \tag{4.3.1}$$

其中 $\varphi_m(x)$是 MMI 部分第 m 阶本征模，c_m 是相应部分的权重系数. 那么，在 z 处的光场为

$$f(x,z)=\sum_m c_m\varphi_m(x)e^{-i\beta_m z}, \tag{4.3.2}$$

式中 β_m 为第 m 阶本征模的传播常数.

为了对上式进行化简，自镜像理论采用了两个假设：一是假定 k_{ey} 远小于 $k_0 n_r (k_{ey}^2+\beta_m^2=k_0^2 n_r^2)$，其中 k_0 为真空中波矢，k_{ey}为波矢的 y 分量；二是认为第 m 阶本征模的有效宽度与 0 阶本征模的有效宽度相等，即 $W_e^m=W_e^0$. 基于以上两个假设，式(4.3.2)可以改写为

$$f(x,z)=\exp(i\beta_0 z)\sum_m c_m\varphi_m(x)\exp\left[i\frac{m(m+2)\pi}{3L_\pi}z\right], \tag{4.3.3}$$

其中 L_π 为拍长，其定义为

$$L_\pi=\frac{\pi}{\beta_0-\beta_1}. \tag{4.3.4}$$

通过仔细分析式(4.3.3)，可知产生输入场的 N 重像的位置为[66]

$$L=\frac{3L_\pi}{N}. \tag{4.3.5}$$

此外，对于输入/输出波导的坐标位置，可以确定如下

$$x_i=\frac{[2i-(N+1)]}{2N}W_e^0. \tag{4.3.6}$$

于是，对于给定宽度的 MMI，可以通过上两式确定 MMI 的长度以及输入/输出波导的坐标位置.

如图 4.12 所示，为 1×4 型 MMI 耦合器光场分布图. 由图可知，在 MMI 耦合器的输出端，出现了 4 重输入场的像，从而实现了功率分配的目的.

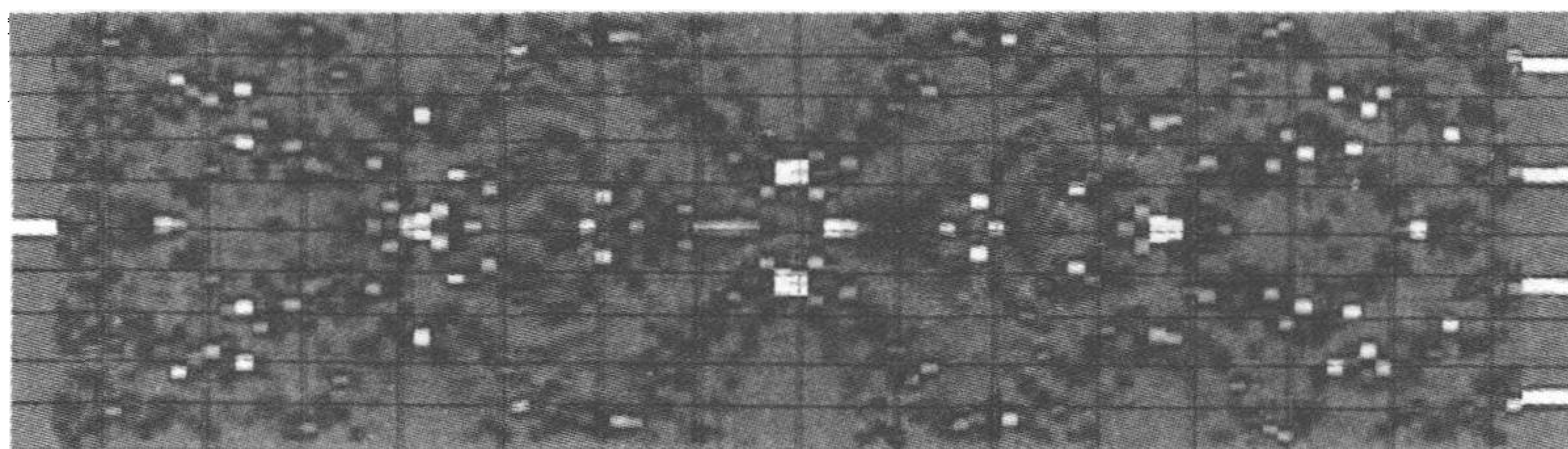


图 4.12 1×4 型 MMI 耦合器光场分布图

4.3.2　多模干涉耦合器的应用

基于上面介绍的自镜像理论，人们纷纷设计了不同的 MMI 器件，以满足实际中的各种需求.

1. 硅纳米线偏振无关 MMI 功分器[67]

由自镜像理论可知，对于 $1\times N$ 型 MMI 耦合器，其第一次 N 重像的位置位于

$$L_{\mathrm{MMI}}=\frac{3L_\pi}{4N}. \tag{4.3.7}$$

为了达到偏振无关的目的，则要求 TE 和 TM 偏振光拍长之差为 0，即 $\Delta L_\pi=L_\pi^{\mathrm{TE}}-L_\pi^{\mathrm{TM}}=0$.

下面以 1×2 型基于 SOI 系统的 MMI 功分器为例，加以说明. 设 SiO_2 的折射率为 1.46，Si 的折射率为 3.46，空气的折射率为 1. 为了防止光场泄漏到硅基底中，设 SiO_2 绝缘层足够厚(比如 2 μm). 选取 $W_{\mathrm{MMI}}=3.86\ \mu\mathrm{m}$，$L_{\mathrm{MMI}}=15.3\ \mu\mathrm{m}$，并利用三维束传播法(3D-BPM)对此结构进行模拟，其光场分布如图 4.13 所示. 对比分析图 4.13(a)和(b)可知，对于 TE 和 TM 偏振光，此结构在相同的位置输出光场几乎相同，从而验证了此器件的偏振无关特性.

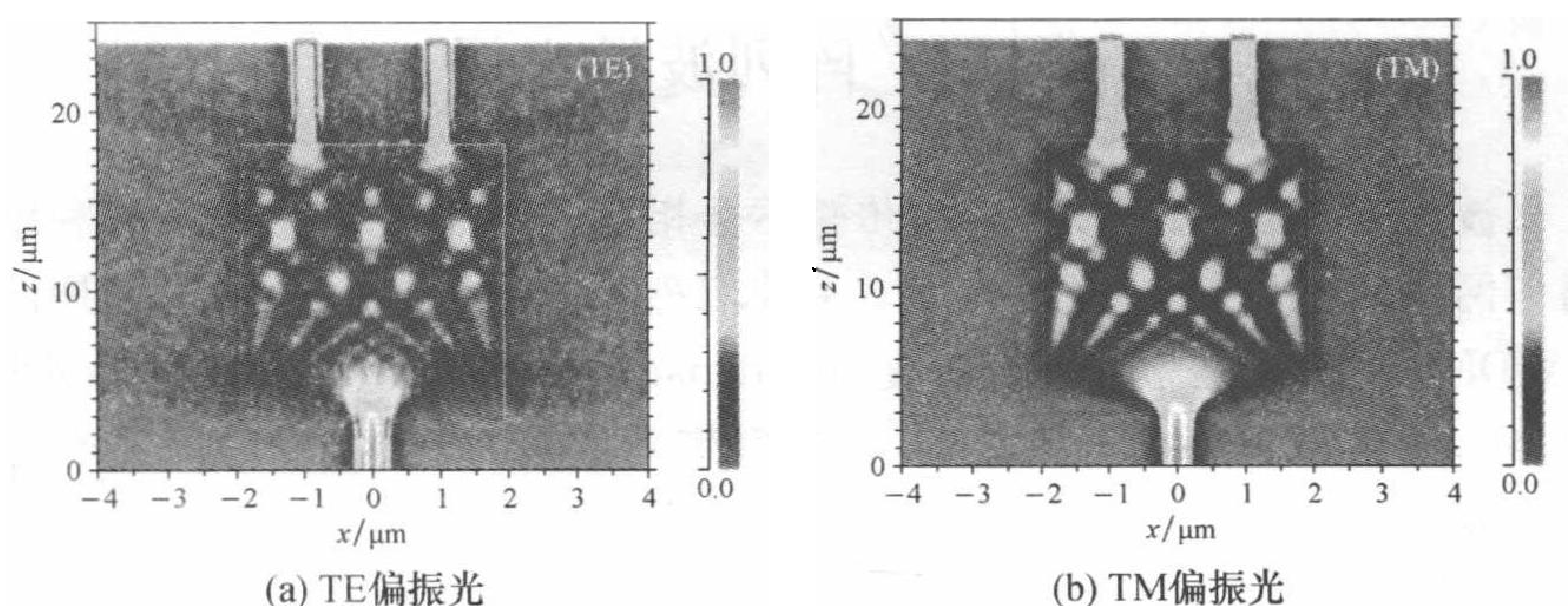


图 4.13　1×2 型基于 SOI 系统 MMI 功分器光场分布图

2. MMI 型 1×2 波长波分复用器[68]

由 SIE 可得：当多模波导长度为 L_π 奇数倍时，多模波导终端得到与输入场相对于多模波导中线呈轴对称的一个反演像；当多模波导长度为 L_π 偶数倍时，多模波导终端得到与输入场完全相同的一个正像. 因此，只要合理选择多模波导的结构参数，使得多模波导的长度为某一波长 L_π 的偶数倍，为另一波长 L_π 的奇数倍，多模干涉耦合器就可以实现两波长分离.

图 4.14 所示为基于 SOI 结构的 1300 nm/1550 nm 的 1×2 波长波分复用器. 1300 nm 和 1550 nm 两个不同波长的光从同一单模波导输入，经多模波导

后，两波长在不同单模波导输出，可实现两波长的分离. 同时，该器件在多模波导的输出端使用了倾斜结构，通过调节倾斜角度，可灵活调节输出波长、功分比等. 理论分析表明，该器件在 1300 nm 和 1550 nm 波长插入损耗分别为 −0.227 dB 和 −0.31 dB，串扰低于 −22 dB，偏振相关损耗分别为 0.033 dB 和 0.01 dB.

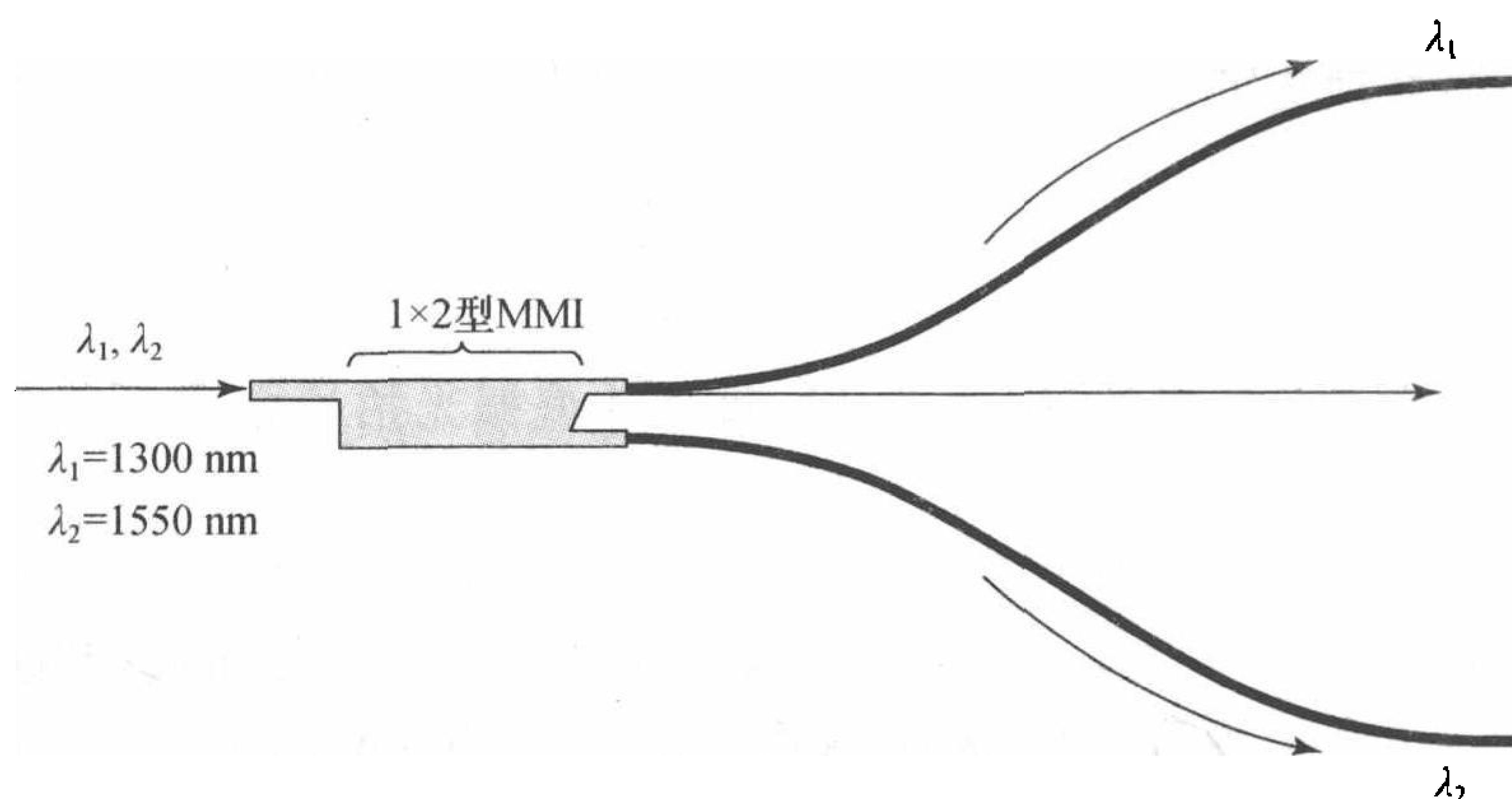


图 4.14 MMI 型 1×2 波长波分复用器

§ 4.4 阵列波导光栅

光波作为载频，在同一通道中传输不会相互干扰，为了充分挖掘和利用现有光通信网络的宽带优势，满足人们对多种信息量呈指数增长的需求，发展波分复用（WDM），尤其是密集波分复用（dense wavelength division multiplex, DWDM）技术，对网络的升级扩容，发展宽带新业务，以及实现超高速通信具有十分重要的意义. 而在 WDM 和 DWDM 系统中，波长复用器和解复用器是关键器件，其性能的优劣对光复用系统的传输质量有决定性影响. 在此背景下，阵列波导光栅（AWG）便孕育而生了.

4.4.1 阵列波导光栅基本概念

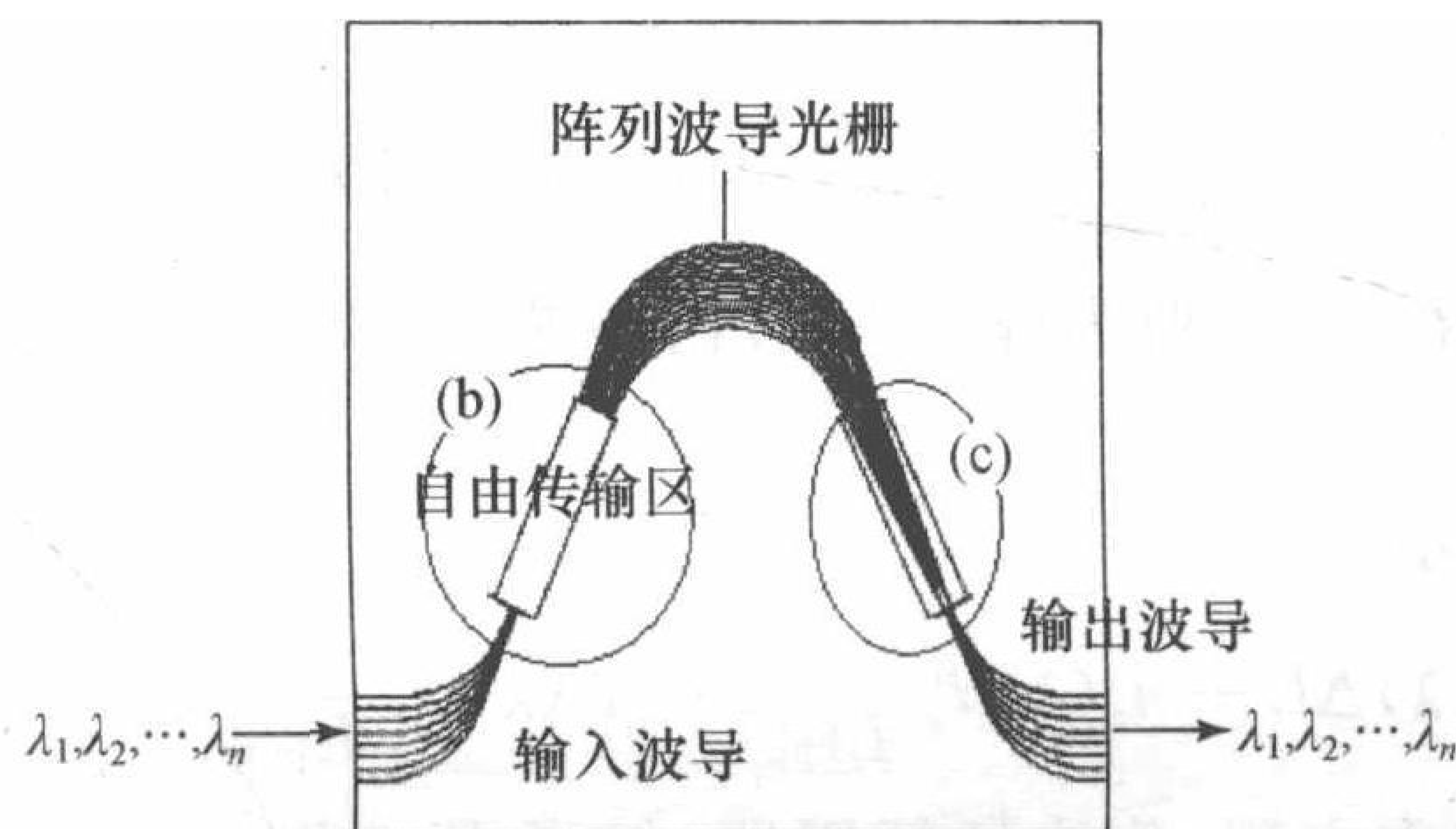
1988 年，Smit 提出了一种新型集成光波导器件——阵列波导光栅[69]. 随后，Vellekoop 和 Smith[70] 报道了第一个工作于短波段的 AWG 器件，而 Takahash 等则[71] 报道了第一个工作于长波窗口的 AWG 器件，Dragone[72] 则将 AWG 的概念从 1×N 型推广到 N×N 型. 简单来说，AWG 是基于干涉原理的波分复用/解复用器件，它是由输入/输出波导阵列、自由传播区平板波导和弯曲波导阵列组成. 各输入/输出波导连接自由传输区的一端以一定的中心间距均匀地排列在罗兰圆的圆周上，弯曲波导阵列中的每条波导正对中心 I/O 波导，均

匀地排列在以中心 I/O 波导为圆心的圆周上. AWG 是第一个将平面波导技术应用于商品化的产品,目前商用流行的是 40 通道 AWG.

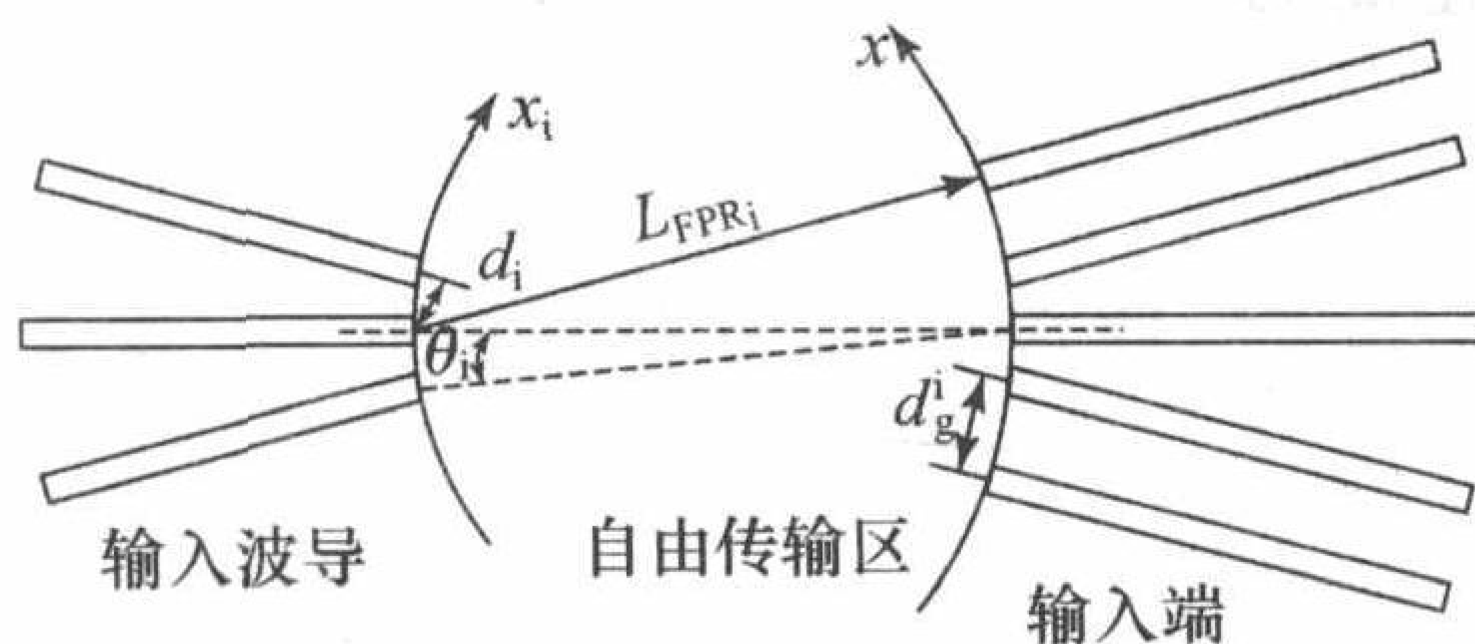
AWG 目前主要发展趋势有两个：一是优化 AWG 的性能,如：减小插入损耗,降低串扰,消除偏振色散,提高中心波长的稳定性和精确度,将通带顶部展宽,提高温度不敏感性等;另一个是拓展 AWG 的功能,利用单片集成技术,构筑系统或网络所需要的功能模块. 日本 NTT 公司在 AWG 的研究方面处于世界领先水平;国内的浙江大学、中科院半导体所、清华大学、华中科技大学、烽火科技集团等相关机构都对 AWG 开展了大量的研究,也取得了比较好的结果.

4.4.2　AWG 设计原理

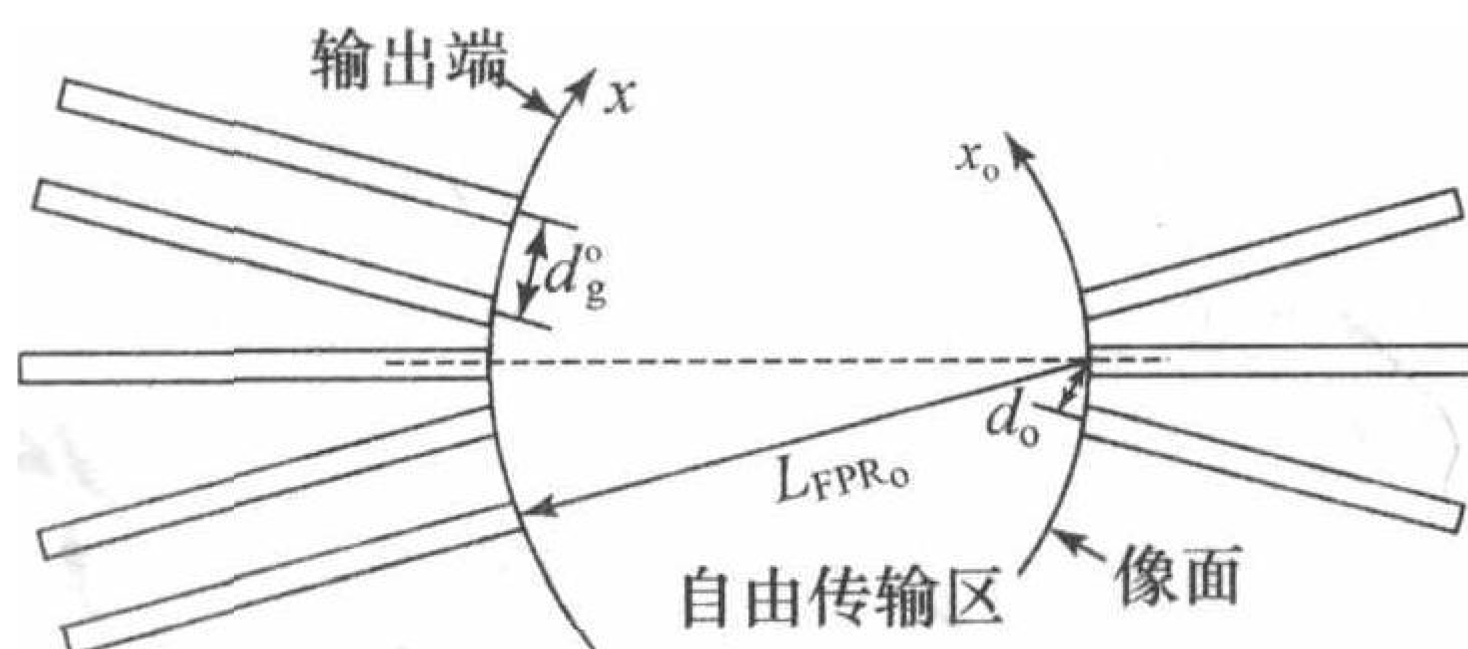
下面以图 4.15 所示 AWG 结构为例[图 4.15(b),(c)所示为图 4.15(a)中椭圆框中的局部放大图],讨论波导光栅的设计原理[52].



(a) AWG结构示意图



(b) 第一自由传输区局部结构图



(c) 第二自由传输区局部结构图

图 4.15　AWG 结构原理图

1. AWG 的光栅方程

AWG 的光栅方程是从光程的角度分析 AWG 的波分复用/解复用功能. 设第 l 条阵列波导长度为

$$L_l = L_o + l\Delta L, \tag{4.4.1}$$

式中 L_o 为最短阵列波导的长度,ΔL 为相邻阵列波导的长度差.

对某一波长 λ,从一波导(x_i)输入,经过 $l, l-1$ 条阵列波导,到达像面上某一点 x_o. 若要在该点干涉加强,则这两条路径的光程差为波长 λ 的 m(m 为衍射级)倍,即[73]

$$\begin{aligned} & n_s(\lambda)\left(L_{\mathrm{FPR_i}} + \frac{d_g^i}{2}\sin\theta_i\right) + n_g(\lambda)[L_o + (l-1)\Delta L] + n_s(\lambda)\left(L_{\mathrm{FPR_o}} + \frac{d_g^o}{2}\sin\theta_i\right) \\ & = n_s(\lambda)\left(L_{\mathrm{FPR_i}} - \frac{d_g^i}{2}\sin\theta_i\right) + n_g(\lambda)(L_o + l\Delta L) \\ & \quad + n_s(\lambda)\left(L_{\mathrm{FPR_o}} - \frac{d_g^o}{2}\sin\theta_i\right) - m\lambda, \end{aligned} \tag{4.4.2}$$

式中 $\sin\theta_i = \dfrac{x_i}{L_{\mathrm{FPR_i}}}$, $\sin\theta_o = \dfrac{x_o}{L_{\mathrm{FPR_o}}}$, $n_g(\lambda)$ 和 $n_s(\lambda)$ 分别为阵列波导和平板波导的有效折射率,$L_{\mathrm{FPR_i}}$ 和 $L_{\mathrm{FPR_o}}$ 分别为输入端自由传播区($\mathrm{FPR_i}$)和输出端自由传播区($\mathrm{FPR_o}$)的长度.

由式(4.4.2)可得

$$n_g(\lambda)\Delta L - n_s(\lambda)d_g^i\frac{x_i}{L_{\mathrm{FPR_i}}} - n_s(\lambda)d_g^o\frac{x_o}{L_{\mathrm{FPR_o}}} = m\lambda, \tag{4.4.3}$$

此式为 AWG 的基本方程. 作为解复用器,仅考虑一条中心波导,即 $x_i = 0$,则式(4.4.3)可以进一步化简为

$$n_g(\lambda)\Delta L - n_s(\lambda)d_g^o\frac{x_o}{L_{\mathrm{FPR_o}}} = m\lambda.$$

当波长 $\lambda = n_g\Delta L/m$ 时,此波长称为中心波长,记做 λ_o. 对于中心波长,式(4.4.3)可化简成

$$\frac{d_g^i}{L_{\mathrm{FPR_i}}}x_i(\lambda_o) = -\frac{d_g^o}{L_{\mathrm{FPR_o}}}x_o(\lambda_o), \tag{4.4.4}$$

即

$$x_o(\lambda_o) = -\frac{R_o}{R_i}\frac{d_g^i}{d_g^o}x_i(\lambda_o), \tag{4.4.5}$$

其中 R_i 为输入波导端罗兰圆直径,R_o 为输出波导端罗兰圆直径.

2. AWG 的色散特性

对于非中心波长,聚焦位置可以用线色散来表示

$$x_o = x_o(\lambda_o) + (\lambda - \lambda_o)\frac{\mathrm{d}x_o}{\mathrm{d}\lambda} = -\frac{L_{\mathrm{FPR_o}}}{L_{\mathrm{FPR_i}}}\frac{d_g^i}{d_g^o}x_i(\lambda_o) + (\lambda - \lambda_o)\frac{\mathrm{d}x_o}{\mathrm{d}\lambda}, \tag{4.4.6}$$

将式(4.4.3)两边对波长微分,可得

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}[n_g(\lambda)]\Delta L-\frac{\mathrm{d}}{\mathrm{d}\lambda}[n_s(\lambda)]d_g^i\frac{x_i}{L_{\mathrm{FPR_i}}}-n_s(\lambda)\frac{d_g^i}{L_{\mathrm{FPR_i}}}\frac{\mathrm{d}x_i}{\mathrm{d}\lambda}$$
$$-\frac{\mathrm{d}}{\mathrm{d}\lambda}[n_s(\lambda)]d_g^o\frac{x_o}{L_{\mathrm{FPR_o}}}-n_s(\lambda)d_g^o\frac{x_o}{L_{\mathrm{FPR_o}}}\frac{\mathrm{d}x_o}{\mathrm{d}\lambda}=m. \tag{4.4.7}$$

为考察像面上聚焦点位置 x_o 与波长 λ 的色散关系 $\mathrm{d}x_o/\mathrm{d}\lambda$,设输入点位置固定(即 $\mathrm{d}x_i/\mathrm{d}\lambda=0$),有

$$\frac{\mathrm{d}x_o}{\mathrm{d}\lambda}=-\frac{1}{n_s(\lambda)}\frac{\mathrm{d}n_s(\lambda)}{\mathrm{d}\lambda}\left(x_i\frac{L_{\mathrm{FPR_o}}}{L_{\mathrm{FPR_i}}}\frac{d_g^i}{d_g^o}\right)-\{n_g(\lambda_o)$$
$$-\lambda_o\frac{\mathrm{d}}{\mathrm{d}\lambda}[n_g(\lambda)]\}\frac{\Delta L}{\lambda_o}\frac{L_{\mathrm{FPR_o}}}{n_s(\lambda)d_g^o}. \tag{4.4.8}$$

将 $x_o=x_o(\lambda_o)+(\lambda-\lambda_o)\frac{\mathrm{d}x_o}{\mathrm{d}\lambda}$代入上式,可得

$$\frac{\mathrm{d}x_o}{\mathrm{d}\lambda}\left[1+\frac{1}{n_s(\lambda)}\frac{\mathrm{d}n_s(\lambda)}{\mathrm{d}\lambda}(\lambda-\lambda_o)\right]=-\frac{1}{n_s(\lambda)}\frac{\mathrm{d}n_s(\lambda)}{\mathrm{d}\lambda}\left[x_i\frac{L_{\mathrm{FPR_o}}}{L_{\mathrm{FPR_i}}}\frac{d_g^i}{d_g^o}+x_o(\lambda_o)\right]$$
$$-\frac{1}{n_s(\lambda)}\left\{n_g(\lambda_o)-\lambda_o\frac{\mathrm{d}}{\mathrm{d}\lambda}[n_g(\lambda)]\right\}\frac{\Delta L}{\lambda_o}\frac{L_{\mathrm{FPR_o}}}{d_g^o}. \tag{4.4.9}$$

忽略小量,式(4.4.9)可近似为

$$\frac{\mathrm{d}x_o}{\mathrm{d}\lambda}=-\frac{N_g(\lambda)\Delta L}{\lambda_o}\frac{L_{\mathrm{FPR_o}}}{n_s(\lambda)d_g^o}. \tag{4.4.10}$$

同理,可得对应于输出波导的线色散为

$$\frac{\mathrm{d}x_i}{\mathrm{d}\lambda}=-\frac{N_g(\lambda)\Delta L}{\lambda_o}\frac{L_{\mathrm{FPR_i}}}{n_s(\lambda)d_g^i}, \tag{4.4.11}$$

式中 $N_g(\lambda)=n_g(\lambda_o)-\lambda_o\frac{\mathrm{d}}{\mathrm{d}\lambda}[n_g(\lambda)]$,$N_g(\lambda)$称为群折射率.

设波分复用通道间隔为 $\Delta\lambda_{ch}$,则相邻输出波导间距 d_o 满足

$$\frac{d_o}{\Delta\lambda_{ch}}=\frac{N_g(\lambda)\Delta L}{\lambda_0}\frac{L_{\mathrm{FPR_o}}}{n_s(\lambda_0)d_g^o}=\frac{N_g(\lambda)}{n_g(\lambda_0)}\frac{L_{\mathrm{FPR_o}}m}{n_s(\lambda)d_g^o}. \tag{4.4.12}$$

同理,相邻输入波导间距 d_i 满足下式:

$$\frac{d_i}{\Delta\lambda_{ch}}=\frac{N_g(\lambda)\Delta L}{\lambda_0}\frac{L_{\mathrm{FPR_i}}}{n_s(\lambda_0)d_i^o}=\frac{N_g(\lambda)}{n_g(\lambda_0)}\frac{L_{\mathrm{FPR_o}}}{n_s(\lambda)d_g^i}, \tag{4.4.13}$$

这里定义色散系数 $D_o=\frac{d_o}{\Delta\lambda_{ch}}$,$D_i=\frac{d_i}{\Delta\lambda_{ch}}$.

若波长 λ 的第 m 阶衍射级位置和波长 $\lambda+\Delta\lambda$ 的第 $m-1$ 阶衍射级位置重合,则称 $\Delta\lambda$ 为自由频谱范围,记为 $\Delta\lambda_{FSR}$.从方程(4.4.2)出发,可得

$$\begin{cases} n_g(\lambda)\Delta L - n_s(\lambda)d_g^i \dfrac{x_i}{L_{FPR_i}} - n_s(\lambda)d_g^o \dfrac{x_o}{L_{FPR_o}} = m\lambda, \\ n_g(\lambda+\Delta\lambda_{FSR})\Delta L - n_s(\lambda+\Delta\lambda_{FSR})d_g^i \dfrac{x_i}{L_{FPR_i}} - n_s(\lambda+\Delta\lambda_{FSR})d_g^o \dfrac{x_o}{L_{FPR_o}} = (m-1)(\lambda+\Delta\lambda_{FSR}). \end{cases} \tag{4.4.14}$$

由此可得

$$\lambda + \left[\Delta L \frac{dn_g(\lambda)}{d\lambda} - d_g^i \frac{x_i}{L_{FPR_i}} \frac{dn_s(\lambda)}{d\lambda} - d_g^o \frac{x_o}{L_{FPR_o}} \frac{dn_s(\lambda)}{d\lambda}\right]\Delta\lambda_{FSR} = (m-1)\Delta\lambda_{FSR}. \tag{4.4.15}$$

忽略 $-d_g^i \dfrac{x_i}{L_{FPR_i}} \dfrac{dn_s(\lambda)}{d\lambda} - d_g^o \dfrac{x_o}{L_{FPR_o}} \dfrac{dn_s(\lambda)}{d\lambda}$，化简得

$$\Delta\lambda_{FSR} = \frac{\lambda}{m\dfrac{N_g(\lambda)}{n_g(\lambda_0)} - 1} \approx \frac{\lambda}{m-1}. \tag{4.4.16}$$

除了自由频谱范围 $\Delta\lambda_{FSR}$ 以外，还有一个常用的物理量——自由空间范围(free spatial range) ΔX_{FSR}，表示的是同一波长相邻衍射级的空间距离.

同理，从方程(4.4.2)出发，可得

$$\begin{cases} n_g(\lambda)\Delta L - n_s(\lambda)d_g^i \dfrac{x_i}{L_{FPR_i}} - n_s(\lambda)d_g^o \dfrac{x_o}{L_{FPR_o}} = m\lambda, \\ n_g(\lambda)\Delta L - n_s(\lambda)d_g^i \dfrac{x_i}{L_{FPR_i}} - n_s(\lambda)d_g^o \dfrac{x_o + \Delta X_{FSR}}{L_{FPR_o}} = (m-1)\lambda. \end{cases} \tag{4.4.17}$$

由此可得

$$\Delta X_{FSR} = \frac{\lambda L_{FPR_o}}{n_s(\lambda)d_g^o}. \tag{4.4.18}$$

由自由频谱范围 $\Delta\lambda_{FSR}$，可得最大通道总数为

$$N_{max} = \frac{\Delta\lambda_{FSR}}{\Delta\lambda_{ch}}, \tag{4.4.19}$$

最大衍射级数 M 为

$$M = \frac{\lambda}{N_{max}\Delta\lambda_{ch}}. \tag{4.4.20}$$

3. AWG 的主要性能指标

(1) 串扰 XT(cross talk).

串扰包括相邻通道串扰 AXT(adjacent cross talk)和非相邻通道串扰. AXT 是指某一输出波导的透过率峰值与相邻带通内功率和的比值，其主要来源是输出波导间的耦合. 利用叠加积分法可以对 AXT 进行估算：

$$\mathrm{AXT} = 20\lg|\eta_c|, \tag{4.4.21}$$

式中 $\eta_c = \dfrac{\int_{-\infty}^{\infty} E_o(x)E_o(x-d_o)\mathrm{d}x}{\int_{-\infty}^{\infty} E_o(x)E_o^*(x)\mathrm{d}x}$，$E_o(x)$ 为输出波导的本征模式.

由上式可知，波导间距增大，输出波导间的耦合减小，则串扰减小.

非相邻通道串扰是指带通范围内信号功率与非相邻带通内功率和的比值.

(2) 频谱的 LdB 带宽，其定义为

$$\Delta\lambda_L = 0.77\frac{\omega_e}{d_o}\Delta\lambda_{ch}\sqrt{L}, \tag{4.4.22}$$

其中 ω_e 为高斯场的 1 dB(或 3 dB)等效带宽. 由式(4.4.22)可知，LdB 带宽和输出波导的间距成反比.

(3) 插入损耗 IL(insertion loss).

它是指输出各通道的最小损耗的最大值

$$\mathrm{IL} = \max_j(\mathrm{IL}_j),\quad \mathrm{IL}_j = \mathrm{IL}(\lambda_{\max,j}). \tag{4.4.23}$$

另外一个与 IL 相关的性能指标是插入损耗均匀性 ILU(insertion loss uniformity)，它是指输出各通道损耗的最大差值

$$\mathrm{ILU} = \max_j(\mathrm{IL}_j) - \min_i(\mathrm{IL}_i). \tag{4.4.24}$$

(4) 回损 RL(return loss).

它是指所有输出状态下输入功率与反射内功率的最大比值

$$\mathrm{RL} = \max_j(\mathrm{RL}_j),\quad \mathrm{RL}_j(\mathrm{dB}) = 10\lg\frac{P_{\mathrm{in},j}}{P_{\mathrm{back},j}}. \tag{4.4.25}$$

(5) 方向性 D(direction).

它定义为输入功率 $P_{\mathrm{in},i}$ 与其他输入口的反射内功率 $P_{\mathrm{back},j}$ 的最大比值：

$$D = \max_j(D_j),\quad D_j(\mathrm{dB}) = 10\lg\frac{P_{\mathrm{in},i}}{P_{\mathrm{back},j}}\ (i \neq j). \tag{4.4.26}$$

4.4.3　AWG 的应用

AWG 除了在(密集)波分复用系统里实现多个通道的合波或分波功能以外，还可和其他器件集成，构成多功能的模块，如波长路由器(wavelength router)、多波长接收器、多波长激光器、波长选择性开关等.

1. 粗波分复用光接收机[74]

图 4.16 所示为日本 NTT 公司研制的以多模输出阵列波导光栅(MM-AWG)为主体的高响应度、通带平坦的粗波分复用(coarse wavelength-division multiplexing，CWDM)光接收机，它能同时覆盖 S，C 和 L 波段. 该器件由 MM-AWG、光电二极管(photodiode，PD)阵列和电路板构成，其中 MM-AWG 相邻

输出信道的间隔为 20 nm，8 信道由 InGaAs PIN-PD 阵列组成，其响应度为 1.1 A/W，封装尺寸为 100 mm×50 mm. 这种 8 信道光接收机可以用于 1.55 μm 波段的 CWDM 系统中，当所有信道同时驱动时，它可以工作在 1.25 Gbit/s 速率下.

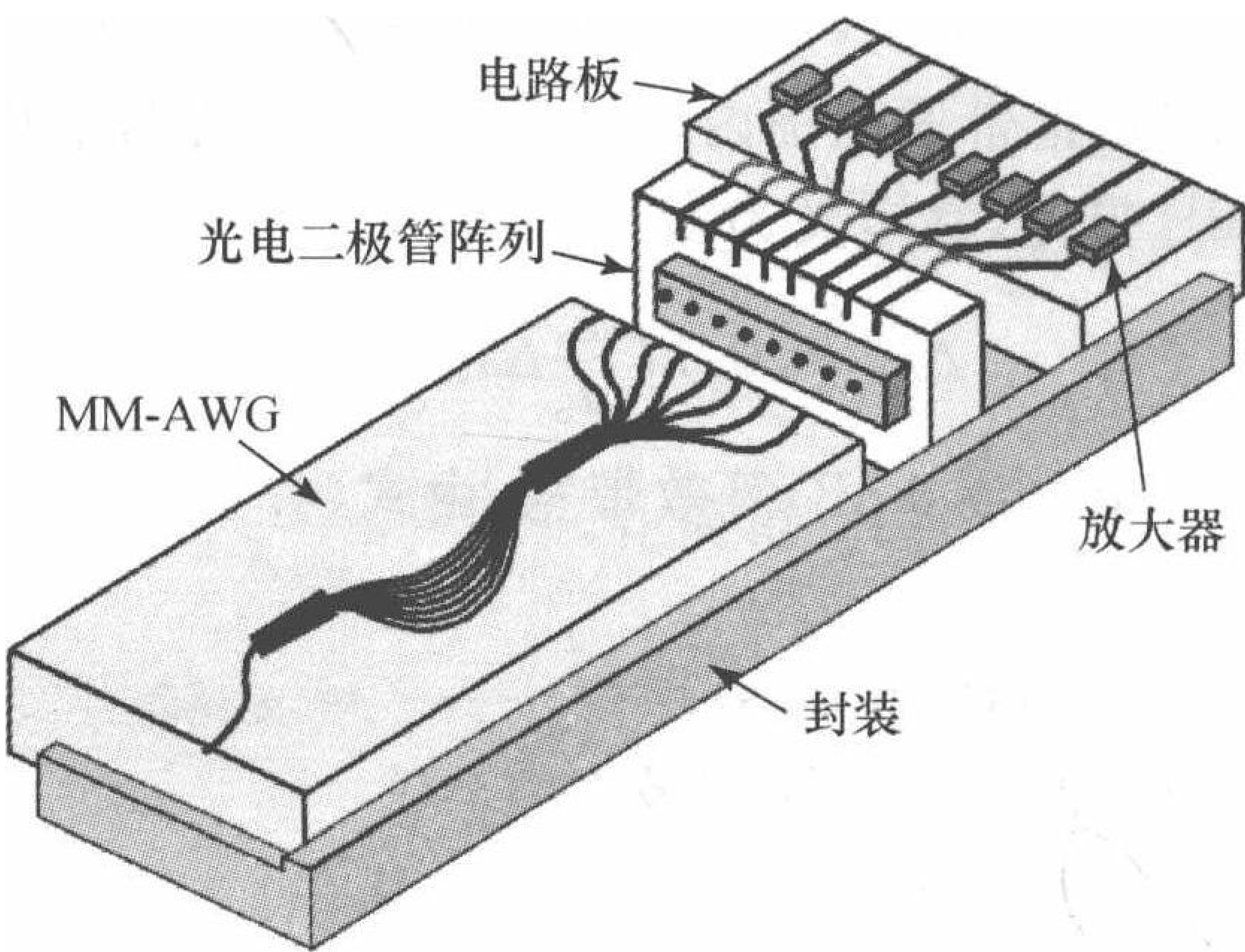


图 4.16 粗波分复用光接收机结构示意图

2. 偏振无关双工器[75]

在设计偏振无关集成光路时，主要的问题在于光波导器件的偏振相关性与光在光纤中传播时偏振状态的不确定性. 为了解决此问题，有课题组基于 AWG 结合二维光栅提出了一种新型的偏振分集集成器件，其结构如图 4.17 所示.[75] 它是由输入光栅、AWG 以及两个输出光栅组成. 其中输入光栅一方面用于将单模光纤的光耦合进平面波导；另一方面还有偏振分束的目的，即将输入光分成相互正交的两束偏振光 P_1 和 P_2. 然后 P_1 和 P_2 分别沿二维光栅不同的波导臂输入 AWG 中，经 AWG 后，P_1 和 P_2 又沿输出光栅的不同波导臂合成、耦合进输

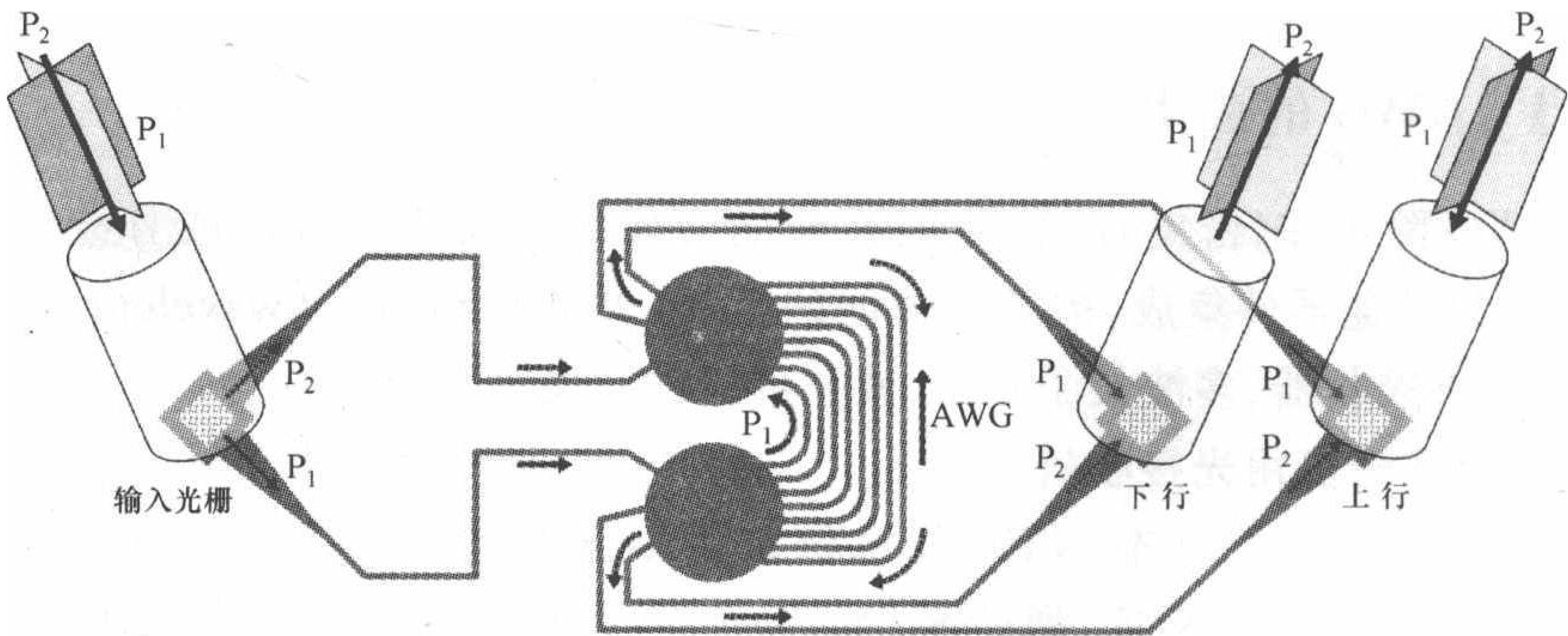


图 4.17 基于 AWG 的偏振无关双工器原理图

出单模光纤中，从而该器件实现了对偏振透明的目的. 实验测得该器件的插入损耗在－6.9～－2.1 dB之间，串扰为－15 dB，偏振相关损耗为0.66 dB.

§4.5　微环谐振腔

4.5.1　微环谐振腔的基本概念

E. A. J. Marcatili于1969年首次提出了微环谐振腔结构[76]，与其他波导结构相比，它特有的环型结构使得满足其谐振条件的光能在微环内循环振荡，故而具有波长选择的功能. 然而，由于当时理论工具、制作工艺和器件材料等方面的原因，对该结构的研究并没有得到很好的重视与发展. 后来，随着时域有限差分法与平面波展开法等方法的发展以及微纳加工工艺的广泛应用，为微环谐振腔结构的实现提供了必要的理论基础和器件制作工艺条件，从此，微环谐振腔的发展进入了一个新的阶段. 1994年，O. Kazuhiro等报道了一个用于光纤通信系统的直径为3.5 mm的环型谐振腔滤波器[77]. 然而，由于该器件尺寸过大，导致谐振腔的自由光谱范围仅为100 GHz，这不能很好地满足实际通信系统的要求. 在此之后，随着微纳加工工艺不断进步以及高折射率差材料体系的应用，波导环腔的尺寸很快降到了微米量级. 目前，已经有微纳尺寸的微环探测器见诸报道[78].

由于微环具有波长选择的功能，因此最初主要用于滤波器中，后来随着对各种有源以及无源材料研究的不断深入，其他性质如非线性[79]也被应用到新的领域中. 目前微环谐振腔主要应用如下几个方面：① 滤波器，如单个微环常被用作滤波器[80]，多个微环级联成阵列时则可被用作复用/解复用器或光开关阵列[81]；② 光开关[82]以及调制器[83]；③ 色散补偿器[84]；④ 波长转换器[85]；⑤ 传感器[86,87].

图4.18所示为一典型双直波导光学环腔结构，该结构由两个平行的直波导以及波导之间的微环组成. 其中a_1为输入端的振幅，a_2和a_3为输出端的振幅. R为微环半径，是圆心与环行波导中心的距离. 光从输入端口输入后，通过直波导和微环之间的耦合进入微环，在环中形成谐振，然后再经过弯曲波导与直波导之间的耦合进入输出端. 由于光波在直波导以及弯曲波导中的传输性质较容易得到，因而在对光学微环谐振腔结构的模型分析中，可以将整个结构看成由左、右两个直-弯波导耦合结构、两段弯曲波导以及四段直波导所构成. 在图4.18左侧的直-弯波导耦合结构中，a_1和b_1分别表示直-弯波导耦合结构的输入端的光波振幅，a_2和b_2分别表示经过直-弯耦合结构后，余留在各波导中的光振幅.

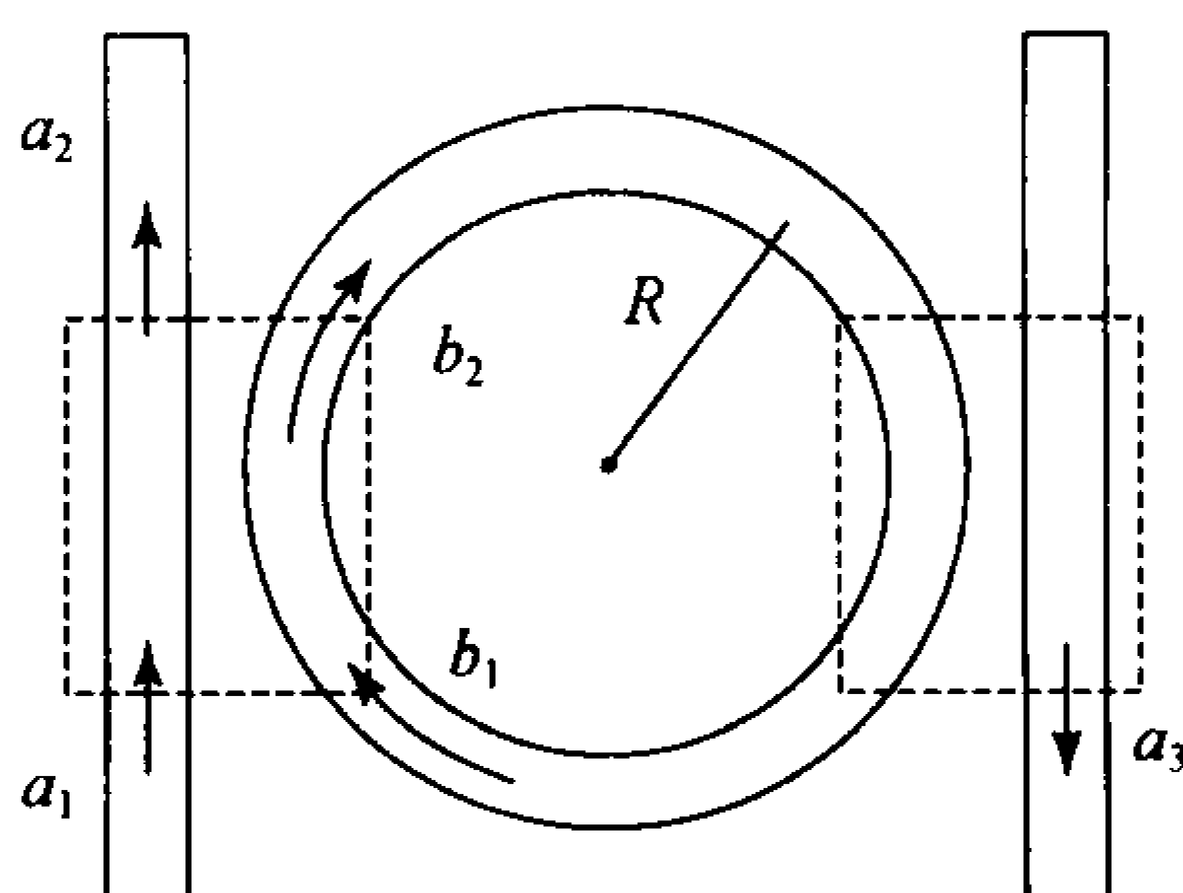


图 4.18 双直波导光学环腔结构图

4.5.2 微环谐振腔模型建立

1. 单直波导微环谐振腔模型

首先，分析单直波导光学环腔结构，如图 4.19 所示，光从直波导的端口 A 输入，经过耦合结构后从端口 B 输出. 该结构中各部分的场强可由以下的关系式表示：

$$b_1 = ta_1 + ka_2, \tag{4.5.1}$$

$$b_2 = -k^* a_1 + t^* a_2. \tag{4.5.2}$$

式(4.5.1)和(4.5.2)中，a_1，a_2，b_1 和 b_2 为各端口的归一化振幅，因此各端口的能量可以由它们的模的平方表示. t，t^*，k 和 k^* 为直-弯波导耦合结构中的各耦合系数，其中 t，t^* 为自耦合系数，k，k^* 为互耦合系数. 由归一化条件，可得

$$|k^2| + |t^2| = 1. \tag{4.5.3}$$

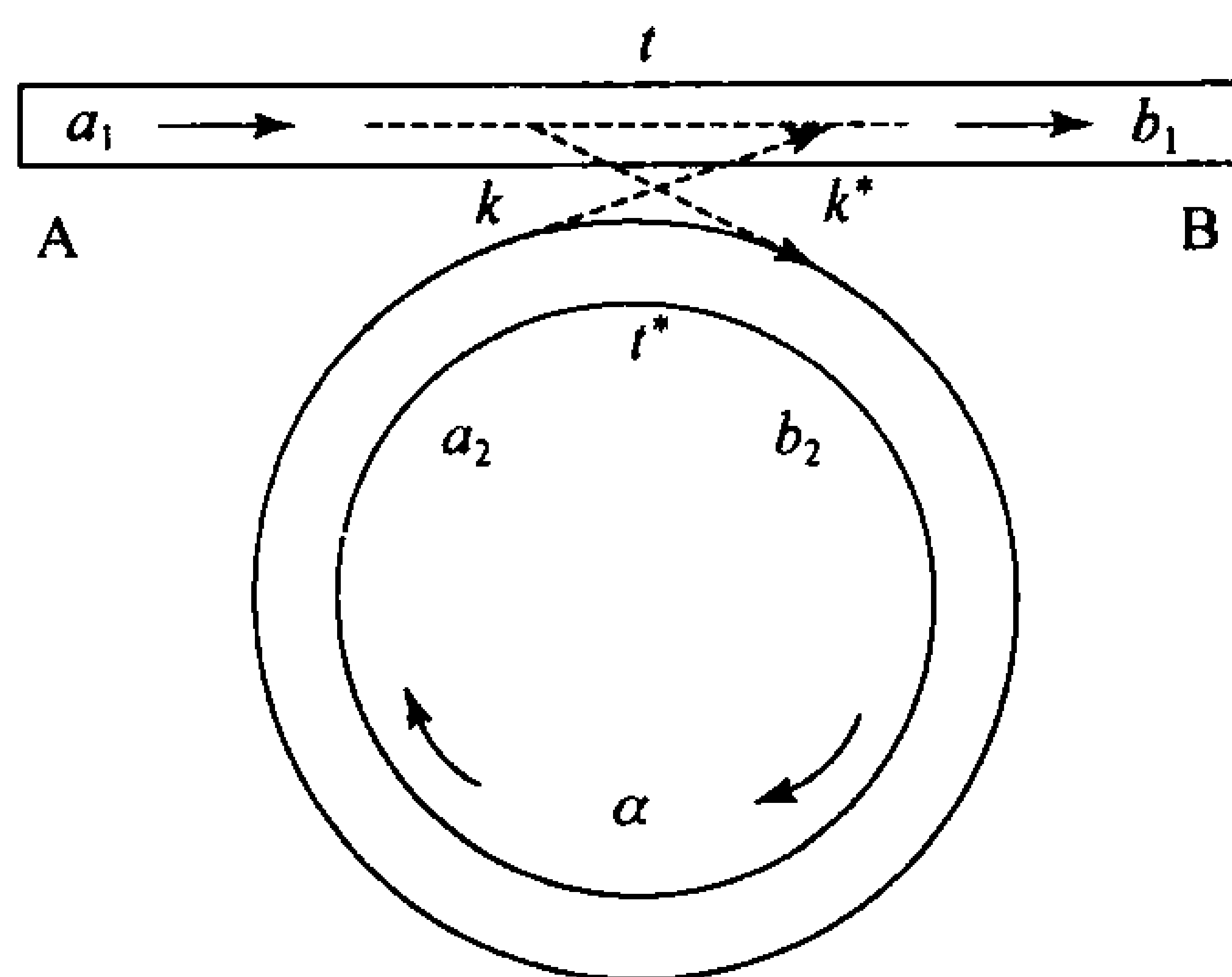


图 4.19 单直波导光学环腔结构理论模型

为了便于运算及归一化，令输入端振幅 $a_1 = 1$. 光在微环中的传输特性可由 a_2 与 b_2 之间的关系，即下式所示

$$a_2 = \alpha e^{i\theta} b_2, \tag{4.5.4}$$

其中 θ 为光在微环内环绕一周的相变，α 为光在微环中传输时的传输系数，为实数. 如光在微环传输时的损耗为 0 时，则 $\alpha=1$. α 可由微环的损耗系数 τ 及周长 L 表示

$$\alpha = \exp(-\tau L). \tag{4.5.5}$$

联立式(4.5.1)，(4.5.2)及(4.5.4)，可得

$$b_1 = \frac{-\alpha + t\mathrm{e}^{-\mathrm{i}\theta}}{-\alpha t^* + \mathrm{e}^{-\mathrm{i}\theta}}, \tag{4.5.6}$$

$$a_2 = \frac{-\alpha k^*}{-\alpha t^* + \mathrm{e}^{-\mathrm{i}\theta}}. \tag{4.5.7}$$

经过微环后，直波导输出端 B 的光强为

$$|b_1|^2 = \frac{\alpha^2 + |t|^2 - 2\alpha|t|\cos(\theta+\phi_t)}{1+\alpha^2|t|^2 - 2\alpha|t|\cos(\theta+\phi_t)}, \tag{4.5.8}$$

其中 $t=|t|\exp(\mathrm{i}\phi_t)$，$\phi_t$ 为直-弯波导耦合区自耦合相变，而在微环中的光强为

$$|a_2|^2 = \frac{\alpha^2(1-|t|^2)}{1-2\alpha|t|\cos(\theta+\phi_t)+\alpha^2|t|^2}. \tag{4.5.9}$$

当微环中的光满足谐振条件时，即 $\theta+\phi_t = m2\pi$，其中 m 为正整数，有

$$|b_1| = \frac{(\alpha-|t|)^2}{(1-\alpha|t|)^2}, \tag{4.5.10}$$

$$|a_2|^2 = \frac{\alpha^2(1-|t|^2)}{(1-\alpha|t|)^2}. \tag{4.5.11}$$

由式(4.5.10)可得，当微环中的传输系数 α 与自耦合系数的绝对值 $|t|$ 相等时，输出端 B 的光强为 0. 此时微环工作在临界耦合状态，如图 4.20 所示，其中取 $\alpha=|t|=0.96$. 临界耦合状态的原理是直波导中的传输光场 ta_1 与微环耦合到直波导中的光场 ka_2 形成相消干涉.

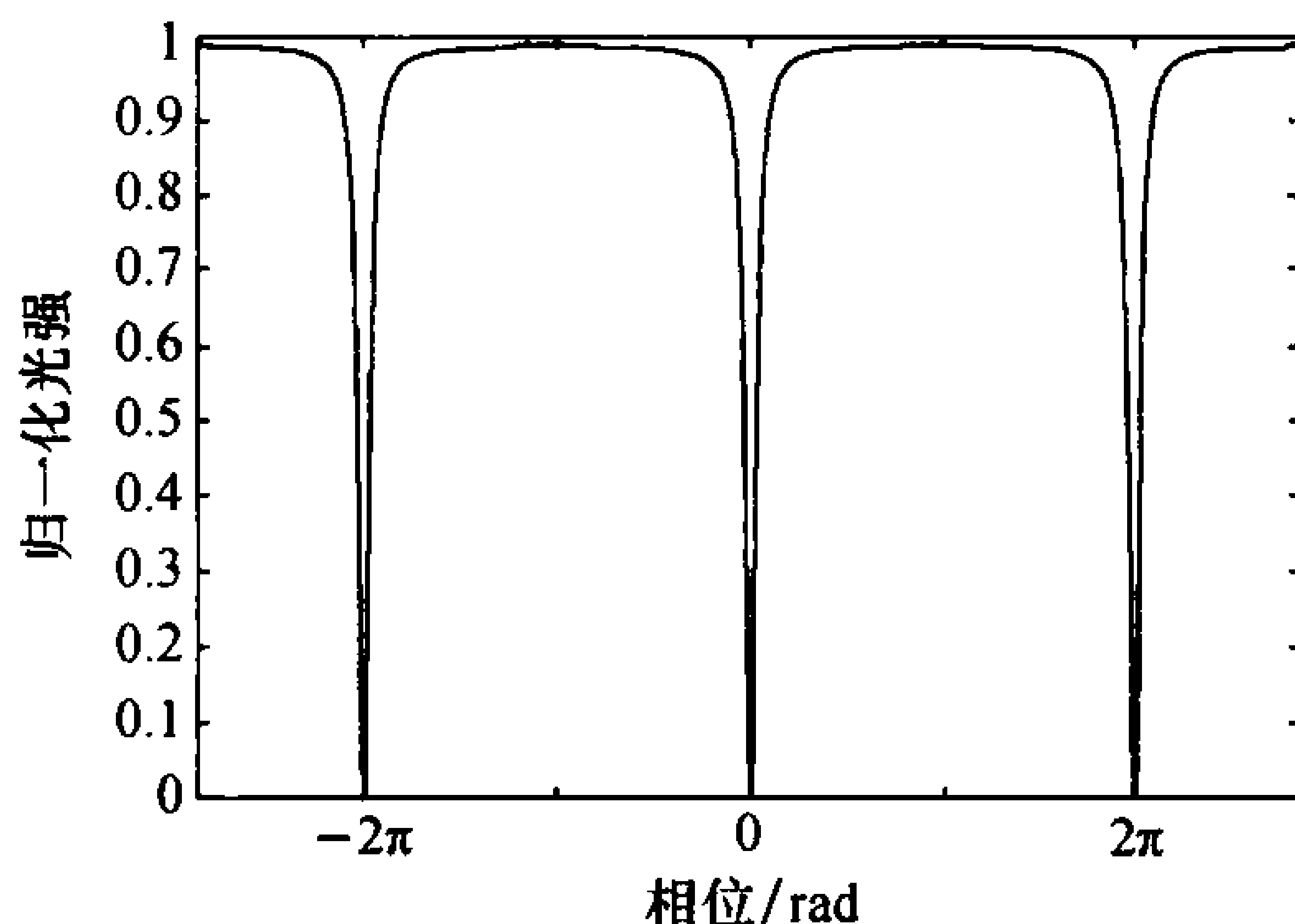


图 4.20　临界耦合状态下单直波导微环的输出曲线

考虑到光在微环中的传输损耗很小，即 $\alpha\approx1$，因此当 $t\to1$ 时，可知在临界耦

合状态下，输出光场为0. 当$|t|$的值在0.7～1之间时，光学微环腔结构的工作状态为弱耦合状态. 若$\alpha\approx1$，则式(4.5.11)可以简化为

$$|a_2|^2 = \frac{1+|t|}{1-|t|}. \tag{4.5.12}$$

从上式可看出，微环中的光强可以大于1. 当$|t|$越趋近于1时，微环中的光强越大，因此光的谐振增强可以得到较高的品质因数Q.

品质因数Q值的定义为中心谐振波长与所对应线宽的比值(或中心频率与频率线宽的比值)，如图4.21所示，其值由式(4.5.13)给出，Q值越高，表明器件对波长的选择性越好，输出峰值越陡峭，因此器件的灵敏度越高.

$$Q = \frac{\lambda_q}{\Delta\lambda_{3\,\mathrm{dB}}}, \tag{4.5.13}$$

其中$\Delta\lambda_{3\,\mathrm{dB}}$为谐振波长3 dB线宽.

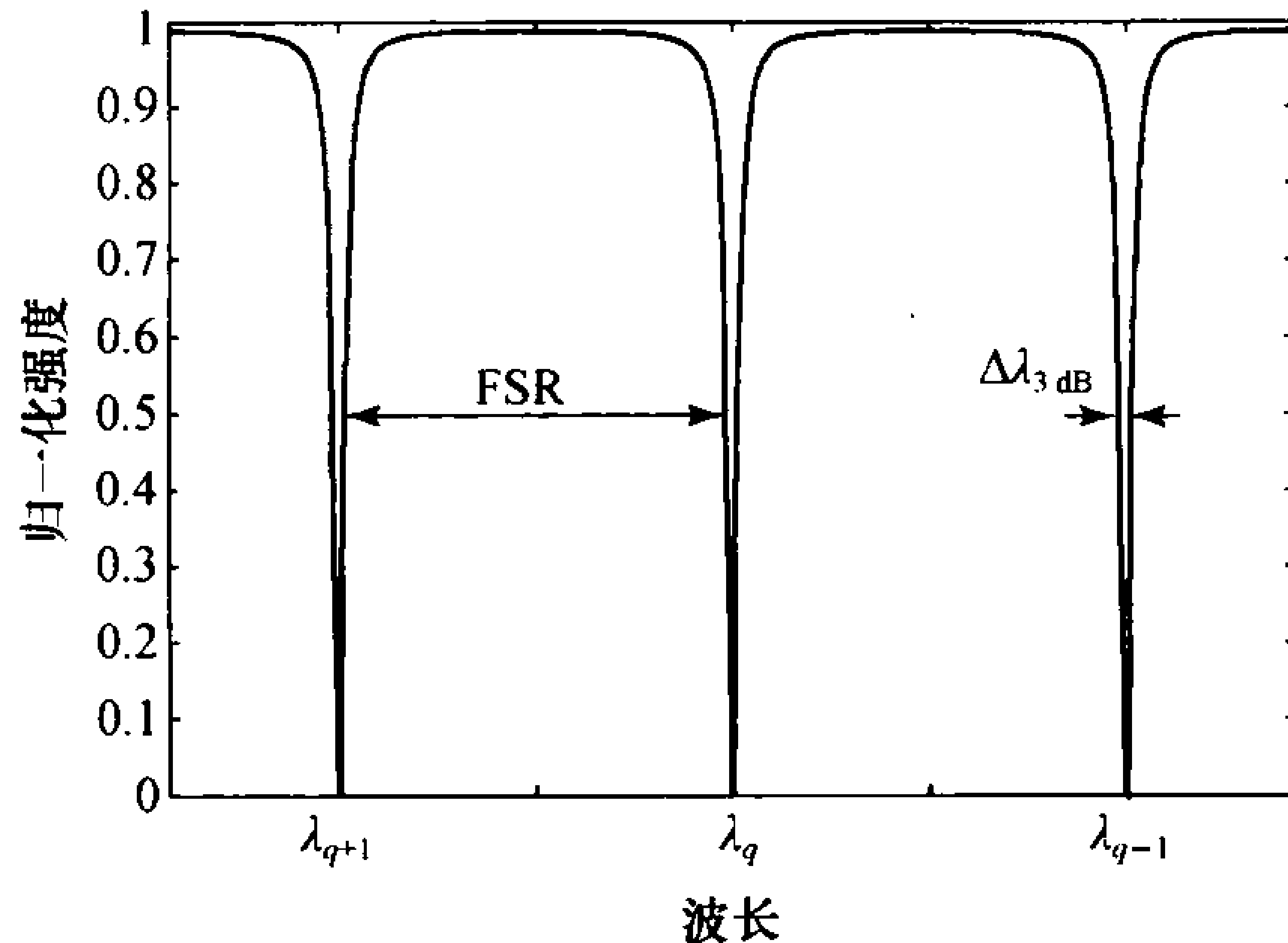


图4.21 单直波导光学环腔透射谱曲线

2. 双直波导微环谐振腔模型

双直波导光学环腔模型如图4.22所示. 各参数定义与前文类似，可以写出各端口振幅之间的关系，其输出端B和D的输出光谱如图4.23所示.

当微环工作在弱耦合状态，$|t|\approx1$时，Q值可由下式给出[88]：

$$Q = \frac{2\pi n_{\mathrm{eff}}}{\lambda_0}\frac{1}{\tau}, \tag{4.5.14}$$

其中n_{eff}为微环的有效折射率.

由上式可知，当损耗系数τ减小时，传输系数接近1，此时可得更高的Q值，其结论与针对单直波导光学环腔计算的结论相吻合. 因此，为了得到更高的Q值，首先需要微环器件工作在弱耦合状态下，同时需要尽量降低光在微环中的传输损耗.

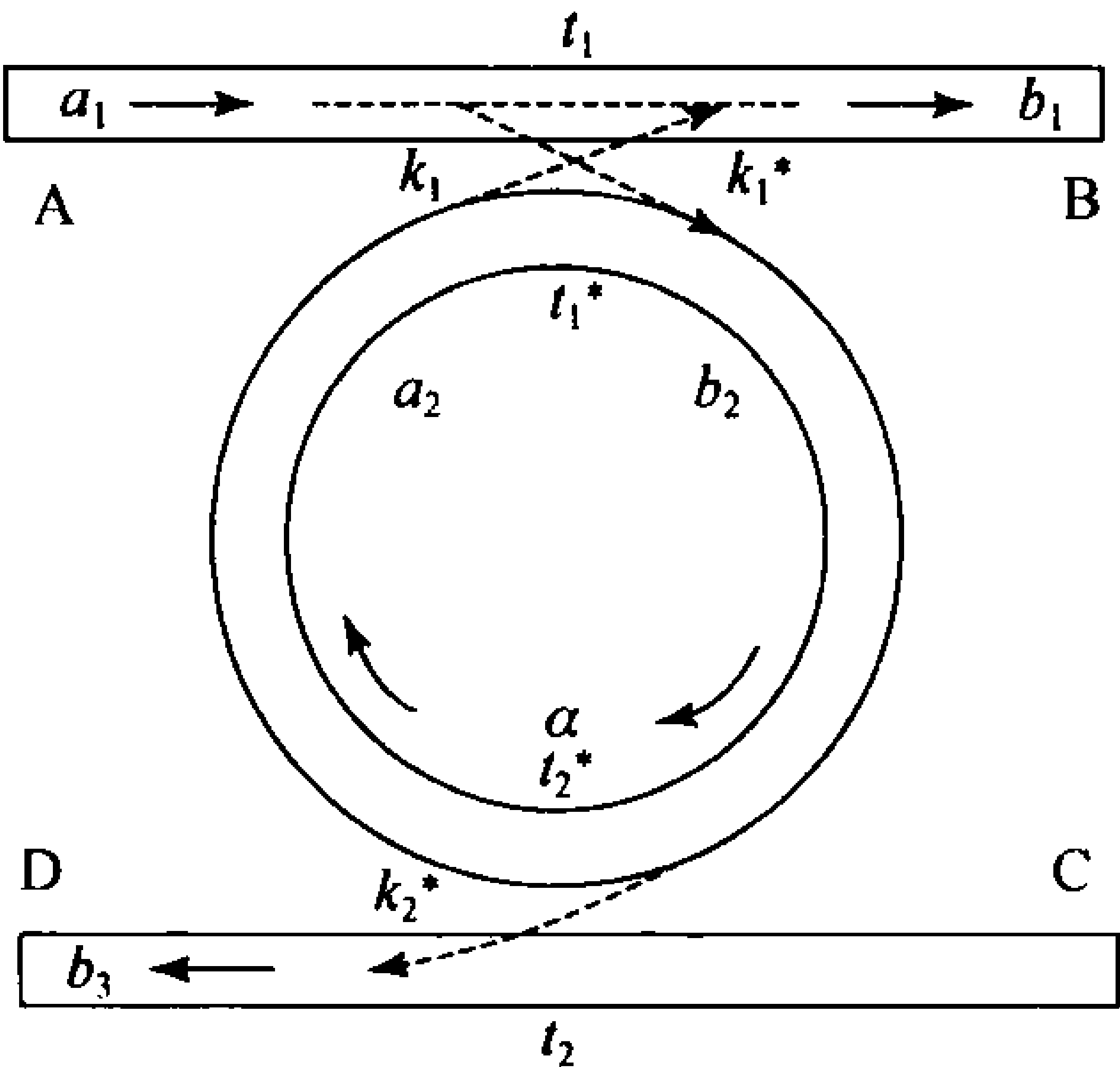


图 4.22 双直波导光学环腔理论模型

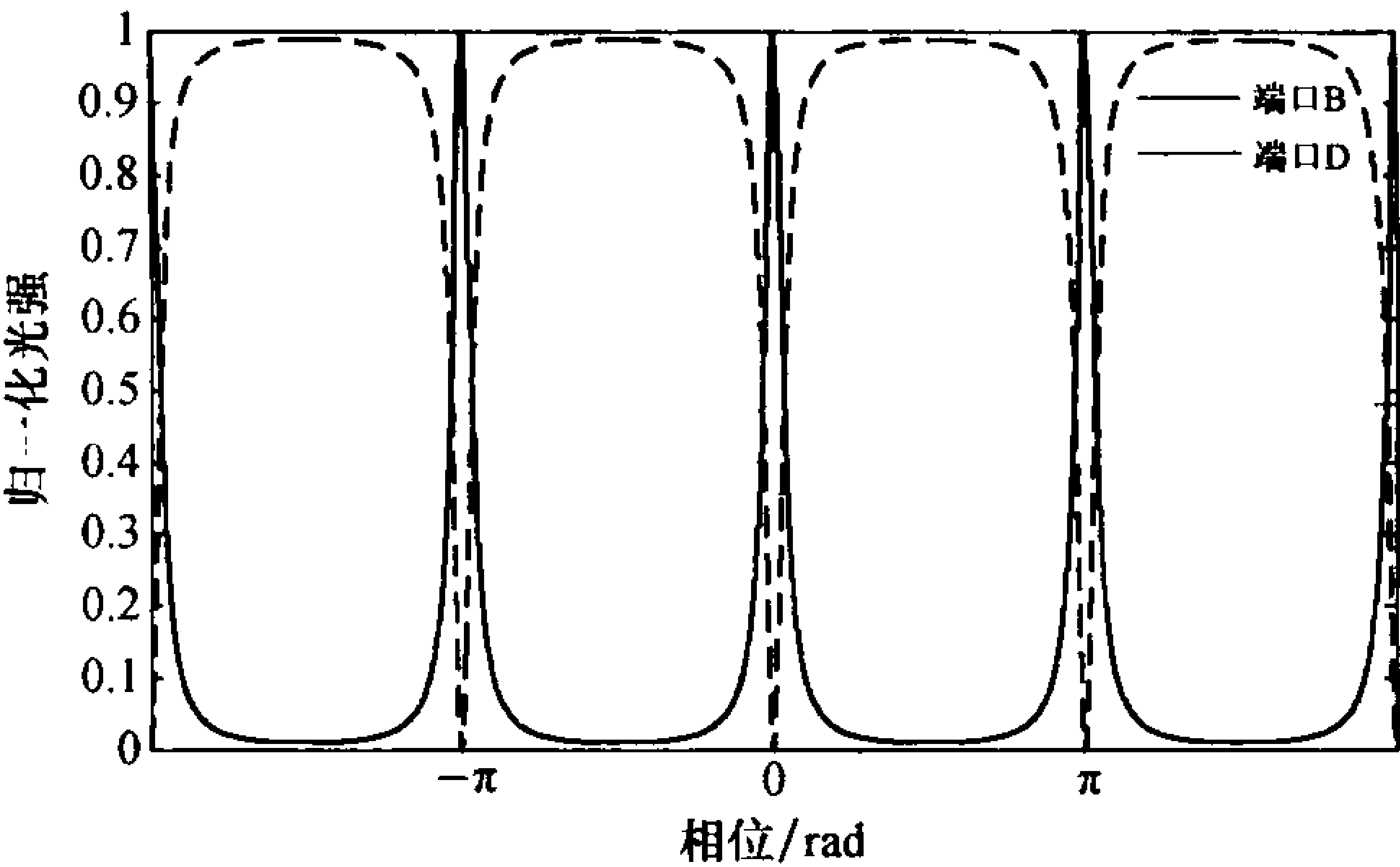


图 4.23 双直波导光学环腔的透射谱曲线

4.5.3 基于微环谐振腔的光集成器件

1. 微环传感器

微纳光学环腔传感芯片，以其特有的光学谐振增强效应而具有极高的灵敏度，且可以实现对微量生物物质的无标记探测. 此外，该传感芯片通常可设计成平面波导结构，可采用成熟的半导体平面工艺加工，易于实现同其他光电芯片和微流控芯片的系统集成，并且容易进行化学表面处理. 另外，谐振效应的引入能够在提高灵敏度的同时大大缩小器件的尺寸，所以基于微环谐振腔的传感器是一种高灵敏度的、可对微量物质进行探测的传感器.

(1) 微环传感器工作原理.

选择如图 4.22 所示双直波导耦合的微环传感器为例进行说明. 在传感器工作时，待分析的样本置于环波导上方区域，如图 4.24 所示. 图中波导层的折射率比上包层和下包层的折射率都较大. 当光的入射角大于布儒斯特角(该角度由波

导层和包层的折射率决定)时,光会在波导层和包层的界面处发生全反射.在这种情况下,大部分光能量都会被限制在波导层内传输.然而,还有一小部分能量会泄漏到波导层之外.这一小部分能量在波导层外以指数形式衰减,称为倏逝场[8].该倏逝场会随着传输条件(如折射率)的变化产生相应的改变.对于环型波导结构而言,待分析样本的加入将改变波导包层的折射率,而该变化能被光波的倏逝场分量所感应,从而引起光波模式有效折射率的改变,并使光在环中的传输特性发生变化,进而影响到波导层中传输的光的性质如光强、相位等.通过检测这些性质的改变,就可以获得被探测物的变化信息.

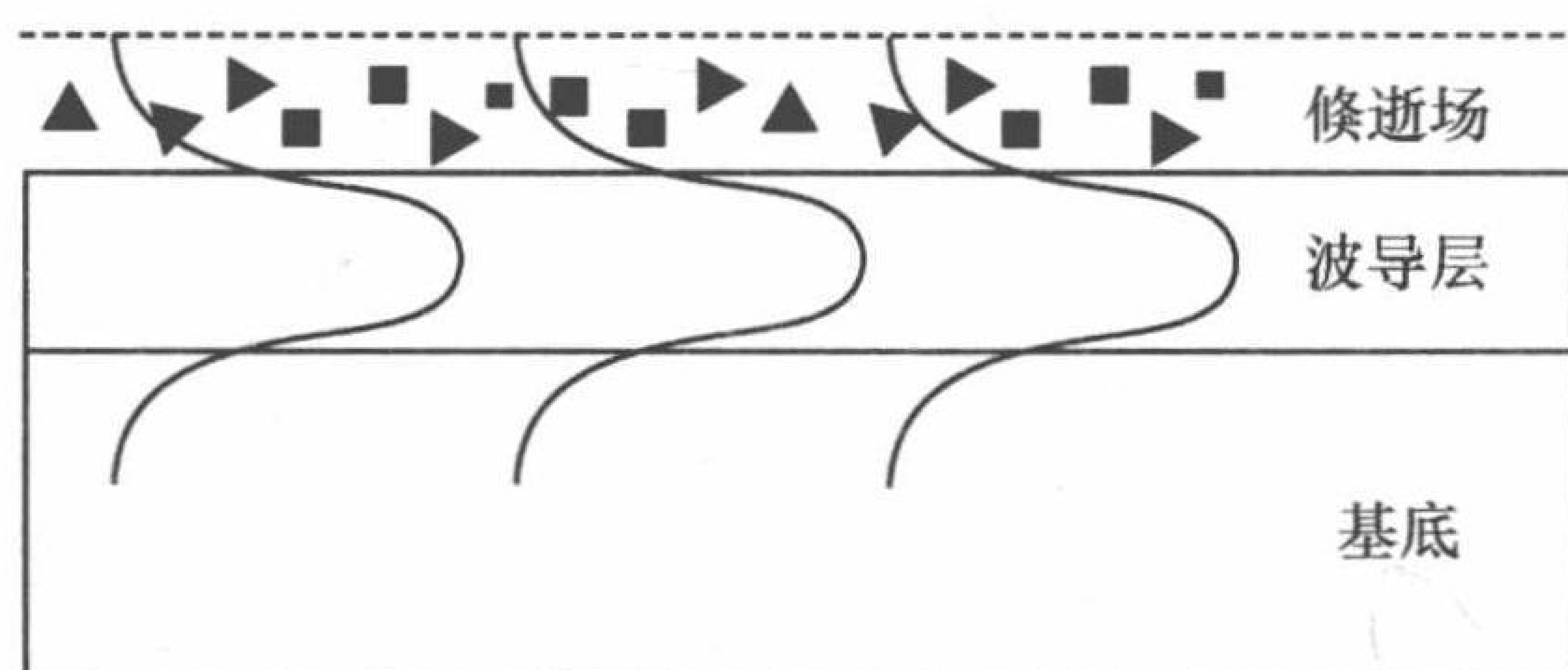


图 4.24 倏逝场传感示意图

根据谐振关系式

$$m\lambda = k_0 n_{\mathrm{eff}} L, \tag{4.5.15}$$

可知,该变化将引起微环响应频谱峰值位置的漂移,如图 4.25 所示.因此,通过监测某一模式下峰值位置的漂移则可建立频移响应曲线,进而得到待测物的相关信息.此外,谐振峰的漂移还会导致在某固定监测波长处的功率响应发生变化,故也可以通过该方法来建立功率响应曲线,进而得到所需的信息.

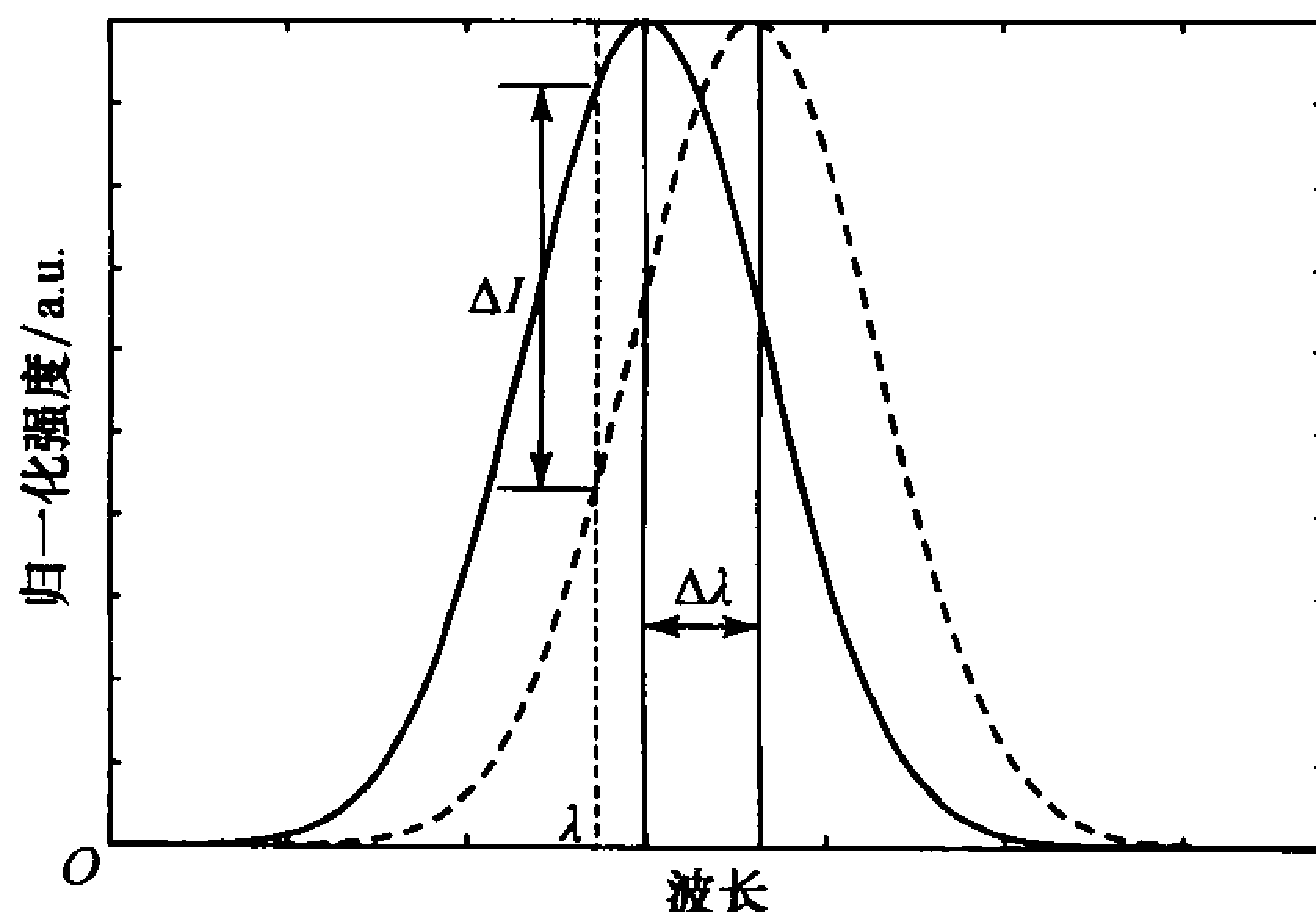


图 4.25 微环传感器工作原理示意图

微环传感器的性能可以通过以下两个参数来衡量:

① 灵敏度.灵敏度所表征的是测量信号随待测信息的变化率,即为响应曲线的斜率.

② 探测极限. 探测极限是传感器探测性能的最终表征参数，因为它直接反映该传感器所能分辨的最小信息变化. 在实际应用中，器件的探测极限除了与核心部件的灵敏度有关之外，还决定于后续探测部件的分辨能力. 对基于倏逝波的集成光波导传感器而言，探测极限通常定义为传感器最小可分辨的包层折射率变化 δn_c，或模式有效折射率变化 δn_{eff}.

(2) 微环传生物感器.

随着生活水平的不断提高，人们对影响身体健康的环境因素和疾病分子的微量检测与监控的要求越来越迫切. 期望通过对这些有害物质和疾病分子的准确检测与监控，及早预防并控制潜在的危害健康的因素，保障人民身体健康. 而微纳光学微环谐振腔传感芯片，具有极高的灵敏度，且可以实现对微量生物物质的无标记探测.

双环谐振腔的干涉作为谐振耦合的实现形式之一，可以实现对微量生物物质的探测. 最近，有研究小组[89]提出了传感环和参考环的双环谐振概念. 在单谐振峰漂移的基础上，当有被测物质加载的情况下，传感环的谐振谱发生漂移，综合考虑传感环和参考环的多谐振峰的干涉作用，通过观察叠加谐振峰位置的漂移，探测微量生物物质. 参考环以及多谐振峰传感的概念对实现新型多功能光学传感具有重要意义.

2. 微环全光信号处理

在以光的 WDM 和 DWDM 技术为基础的现代光通信系统中，网络的各个节点要完成光-电-光的转换，仍以电信号的信息处理速度进行交换，而其中的电子器件在适应高速、大容量的需求上，存在着诸如带宽限制、时钟偏移、严重串扰、高功耗等缺点，由此产生了通信系统中的“电子瓶颈”现象. 因此，实现全光信号处理，发展全光网络，对于消除“电子瓶颈”现象具有十分重要的意义. 全光信号处理是一个总的概念，其包括：波长转换、码型转换、逻辑门、光波分复用/解复用等. 全光网络以其良好的透明性、波长路由特性、兼容性和可扩展性，已成为下一代高速宽带网络的首选.

根据微环结构的谐振增强效应，可将其用于全光信号处理的功能器件中.

(1) 微环全光波长转换器.

在光通信中，波长路由节点将输入波长变换到不同的输出波长上称为波长转换. 目前，波长转换有很多种实现方案，而其中四波混频(four wave mixing, FWM)型波长转换保留了原始信号的相位和幅度信息，是严格透明性的波长转换方案，因此引起了人们的关注.

图 4.26 为基于硅纳米线波导的微环全光波长转换器 SEM 照片[90]，经测试其 FSR 和 Q 值分别为 8 nm 和 23 000. 该结构的反应时间约为 100 ps，这使得其能够应用于高速、宽带的光通信系统中. 此外，仅需几毫瓦的峰值泵浦功率就能

驱动该器件；当利用反偏压雪崩二极管泵浦时，其转换效率接近 0 dB. 此类微环波长转换方案的提出为片上全光通信及信号处理提供了一种可行的途径.

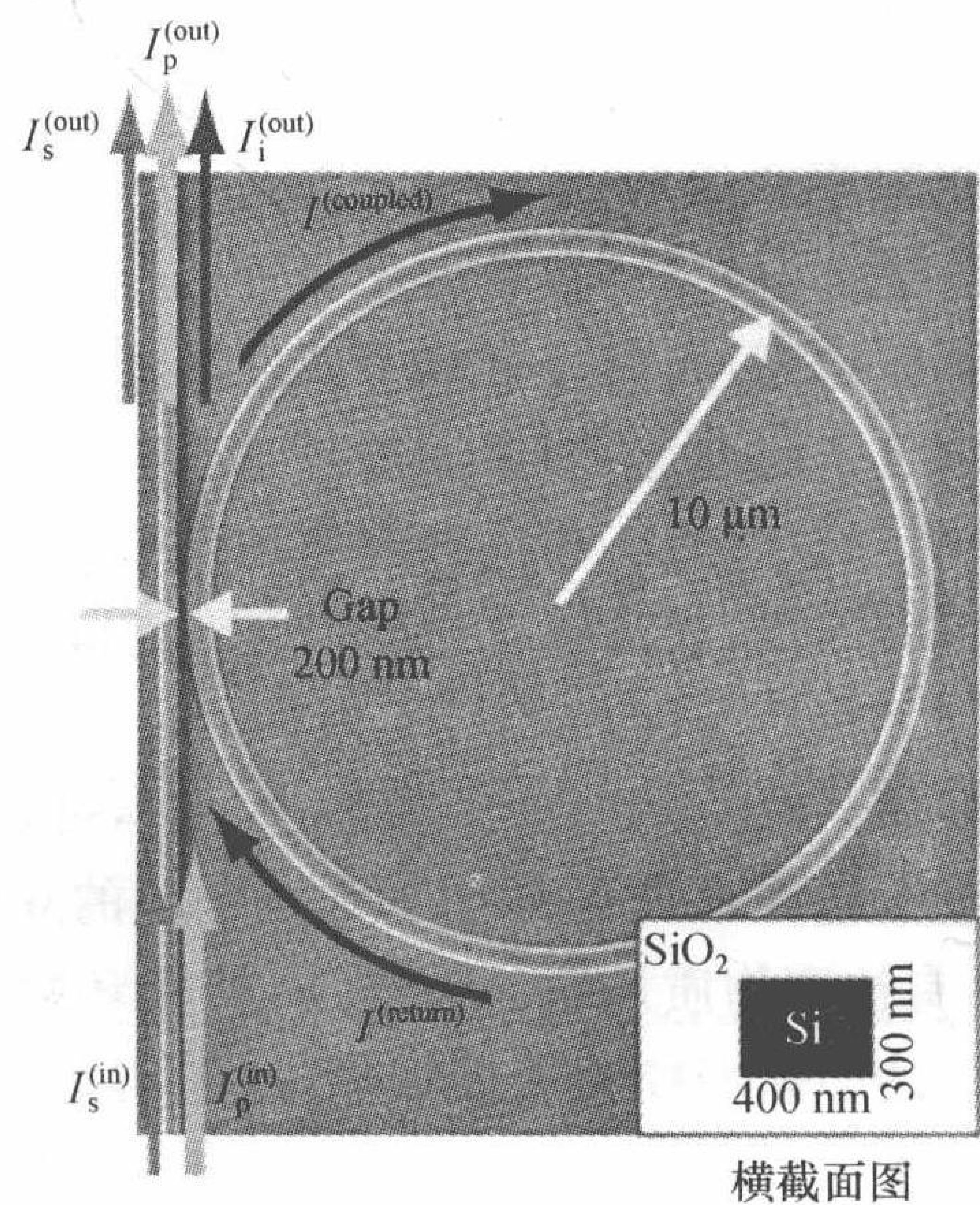


图 4.26 硅纳米线波导微环 SEM 照片

在图 4.27 中，展示了一个 FWM 频谱的例子. 将泵浦光调谐为一个谐振波长 1550 nm，输入光调谐为另一个谐振波长 1544 nm，由于微环谐振的作用，可观察到两个转换频谱（C_1 和 C_2），其中靠近泵浦光的 C_1 为所期望的转换波长.

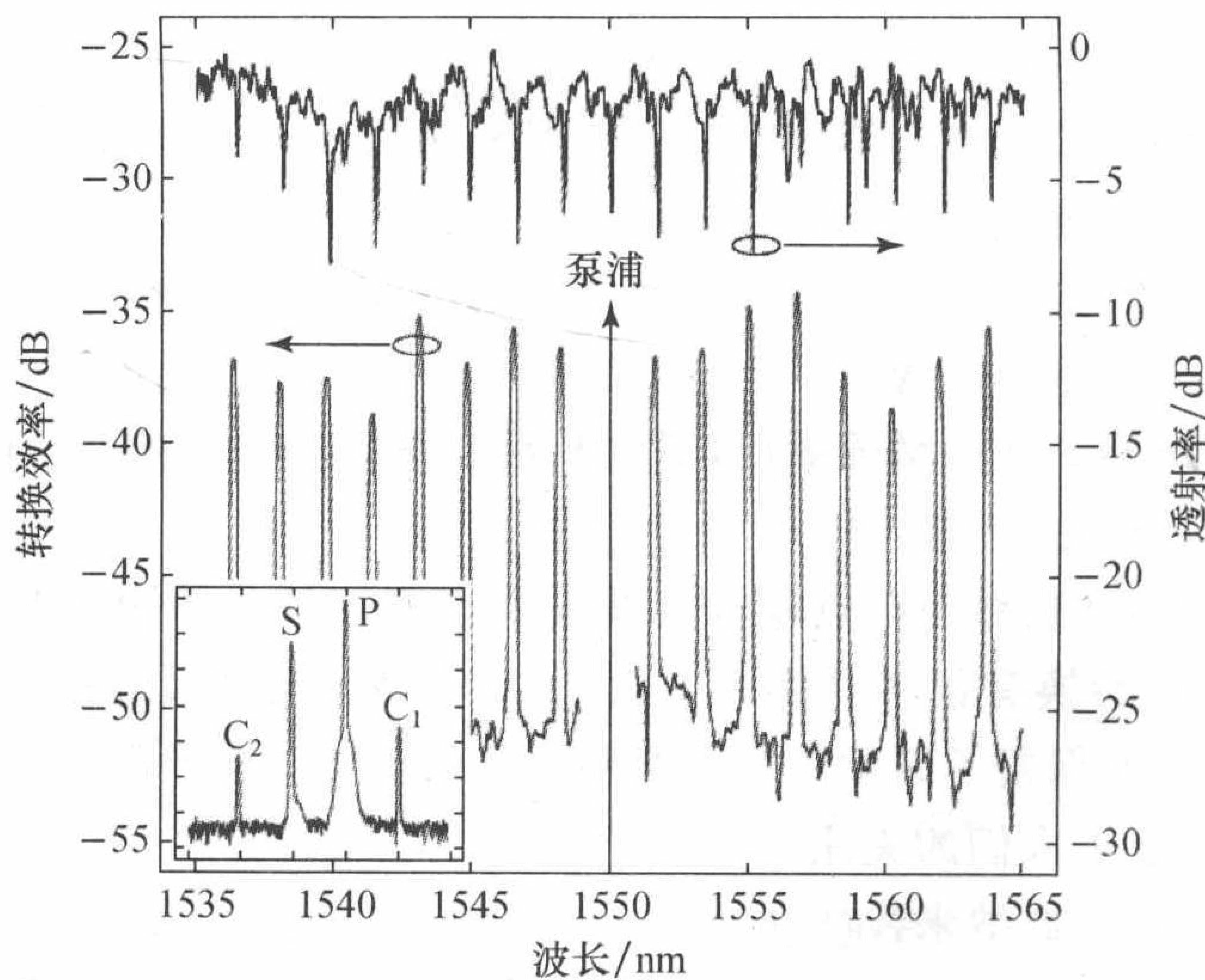


图 4.27 波长与测试转换效率及透射率的关系，图中的小插图 P 表示 FWM 的泵浦光，S 为信号光，C_1 为期望转换光，C_2 为附加转换光

(2) 微环全光码形转换.

利用级联微环的谐振作用,还可实现全光码形转换.

图 4.28 所示为基于级联微环谐振腔的从非归零码(non-return-to-zero, NRZ)到二进制移相键控码(binary phase-shift-keying, BPSK)的码形转换系统示意图[91],其工作原理如下所述:

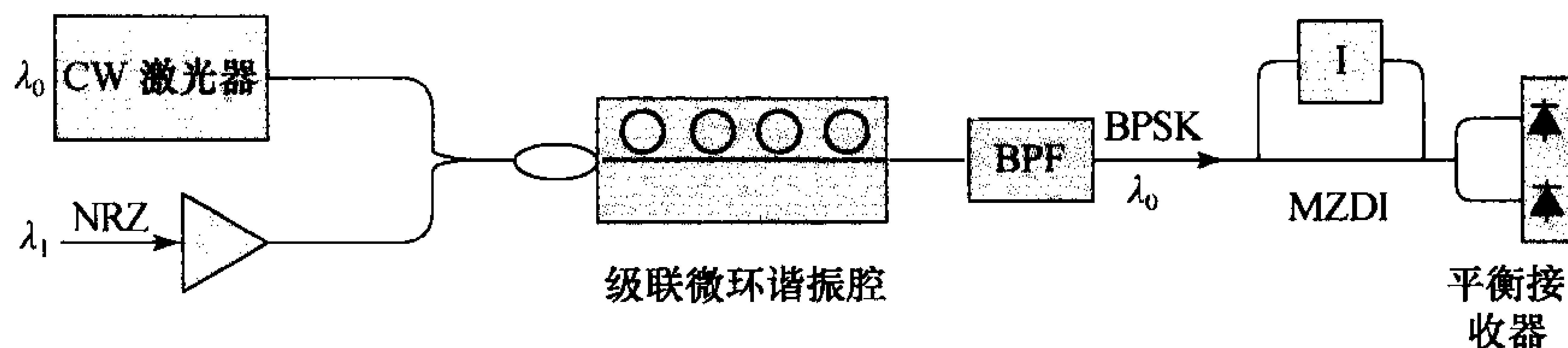


图 4.28　NRZ/BPSK 码形转换系统框图

波长为 λ_0 的弱连续探测光与波长为 λ_1 的 NRZ 泵浦光同时耦合进级联微环谐振腔. 为了使得耦合进微环的光强最大,设 λ_0 和 λ_1 都为谐振波长. 由于泵浦光交叉相位调制(cross-phase modulation)的作用,探测光相位在传输过程中会发生改变,从而使得其相位偏离谐振条件,造成了很大的有效传输相移. 通过正确设置工作条件,可以使得当泵浦光出现时(输入为"1"),探测光经过级联微环谐振腔后,产生了 π 相移;当泵浦光不出现时(输入为"0"),探测光不产生相移. 通过正确设置泵浦光功率,可以使得探测光幅度几乎保持不变,因为在接近谐振位置时,探测光幅度的改变相对于相位改变来说几乎可以忽略不计[92]. 通过该过程便能实现 NRZ 与 BPSK 的码形转换. 光学带通滤波器(bandpass filter, BPF)能够阻止波长为 λ_1 的 NRZ 码,而通过波长为 λ_0 的 BPSK 码. 经过 M-Z 延迟干涉仪(Mach-Zehnder delay interferometer, MZDI) 1 位时间的延迟,BPSK 码可以被解调出来,利用常规的平衡接收器可以接收解调的 BPSK 码.

§4.6　总　　结

光无源器件是现代光学技术中非常重要的一类器件,而硅基光子学无源器件因尺寸小、损耗低、易于与其他器件集成等优点而成为最具应用潜力的器件之一. 本章对光栅器件、光子晶体平板波导器件、多模干涉耦合器、阵列波导光栅器件以及微环谐振腔给出了概略性的介绍,并给出了一些具体应用实例. 但必须指出, 本章只是给出硅基光子学无源器件研究中很小一部分内容的启发性介绍,更多内容读者可以通过查阅文献获取.

参考文献

[1] Suhara T and Nishihara H. Integrated optics components and devices using periodic structures. IEEE J. Quant. Electr., 1986, 22(6): 845.

[2] Day I, Evans I, Knights A, et al. Tapered silicon waveguides for low insertion loss highly-efficient high-speed electronic variable optical attenuators. Proc. Opt. Fiber Communications Conference, Technical Digest Series, 2003.

[3] Sure A, Dillon T, Murakowski J, et al. Fabrication and characterization of three-dimensional silicon tapers. Opt. Exp., 2003, 11:3555.

[4] Pavesi L, Lockwood D J. Silicon Photonics. Berlin: Springer, 2004.

[5] Dai D, He S, Tsang H K. Bilevel mode converter between a silicon nanowire waveguide and a larger waveguide. J. Lightwave Technology, 2006, 24:2428.

[6] Almeida V R, Panepucci R R, Lipson M. Nanotaper for compact mode conversion. Opt. Lett., 2003, 28:1302.

[7] Janz S, Lamontagne B, Delage A, et al. Single layer a-Si GRIN waveguide coupler with lithographically defined facets. 2nd Int. Proc. Conf. on Group IV Photonics, 2005.

[8] Lu Z, Prather D W. TIR-Evanescent coupler for fiber to waveguide integration of planar optoelectronic devices. Opt. Lett., 2004, 29: 1784.

[9] Taillaert D, Bogaerts W, Bienstman P, et al. An out-of-plane grating coupler for efficient butt-coupling between compact planar waveguides and single-mode fibers. IEEE J. Quant. Electr., 2002, 38:949.

[10] Butler J K, Sun N H, Evans G A, et al. Grating-assisted coupling of light between semiconductor and glass waveguides. J. Lightwave Technology, 1998, 16: 1038.

[11] Orobtchouk R, Schnell N, Benyattou T, et al. New arrow optical coupler for optical interconnect. Proc. IEEE International Conference on Interconnect Technology, 2003.

[12] Masanovic G Z, Passaro V M N, Reed G T. Dual grating-assisted directional coupling between fibers and thin semiconductor waveguides. IEEE Photo. Tech. Lett., 2003, 15(10): 1395.

[13] Ang T W, Reed G T, Vonsovici A, et al. Effects of grating heights on highly efficient unibond SOI waveguide grating couplers. IEEE Photo. Tech. Lett., 2000, 12: 59.

[14] Wang B, Jiang J, Nordin G. Compact slated grating couplers. Opt. Exp., 2004, 12(15): 3313.

[15] Moharam M G, Grann E B, Pommet D A. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. J. Opt. Soc. Am. A, 1995, 12(5): 1068.

[16] Yee K S. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. IEEE Trans. Antennas Propag., 1966, 14: 302.

[17] Dakss M L, Kuhn L, Heidrich P F, et al. Grating coupler for efficient excitation of optical guided waves in thin films. Appl. Phys. Lett., 1970, 16: 523.

[18] Wang B, Jiang J, and Nordin G. Embedded slanted grating for vertical coupling between fibers and silicon-on-insulator planar waveguides. IEEE Photo. Tech. Lett., 2005, 17 (9): 1884.

[19] Feng J and Zhou Z. Polarization beam splitter using a binary blazed grating coupler. Opt. Lett., 2007, 32: 1662.

[20] Spuhler G L, Weingarten K J, Grange R, et al. Semiconductor saturable absorber mirror structures with low saturation fluence. Appl. Phys. B, 2005, 81(1): 27.

[21] Liu Z, Lin P-T and Wessels B W. Cascaded Bragg reflectors for a barium titanate thin film electro-optic modulator. J. Opt. A: Pure Appl. Opt., 2008, 10(1): 5302.

[22] Huang M C Y, Zhou Y and Chang H C J. A surface-emitting laser incorporating a high-index-contrast subwavelength grating. Nature Photo., 2007, 1:119.

[23] Chung I S, Mork J, Gilet P, et al. Subwavelength grating-mirror VCSEL with a thin oxide gap. IEEE Photo. Tech. Lett., 2008, 20(2): 105.

[24] Quack N, Blunier S, Dual J, et al. Tunable resonant cavity enhanced detectors using vertically actuated MEMS mirrors. J. Opt. A: Pure Appl. Opt., 2008, 10(4): 4015.

[25] Cheben P, Janz S, Xu D, et al. A broad-band waveguide grating couplers with a subwavelength grating mirror. IEEE Photo. Tech. Lett., 2006, 18(1):13.

[26] Mateus C F R, Huang M C Y, L Chen, et al. Broad-band mirror (1.12~1.62 μm) using a subwavelength grating. IEEE Photo. Tech. Lett., 2004, 16(7):1676.

[27] Liu Z S and Magnusson R. Concept of multiorder multimode resonant optical filters. IEEE Photo. Tech. Lett., 2002, 14(8): 1091.

[28] Liu Z S, Tibuleac S, Shin D, et al. High-efficiency guided-mode resonance filter. Opt. Lett., 1998, 23(19): 1556.

[29] Mateus C F R, Huang M C Y, Deng Y, et al. Ultrabroadband mirror using low-index cladded subwavelength grating. IEEE Photo. Tech. Lett., 2004, 16(2): 518.

[30] Chen L, Huang M C Y, Mateus C F R, et al. Fabrication and design of an integrable subwavelength ultrabroadband dielectric mirror. Appl. Phys. Lett., 2006, 88 (3): 1102.

[31] Wu H, Mo W, Hou J, et al. A high performance polarization independent reflector based on a multilayered configuration grating structure. J. Opt., 2010, 12(4): 5703.

[32] Yablonovitch E. Inhibited spontaneous emission in solid-state physics and electronics. Phys. Rev. Lett., 1987, 58 (20): 2059.

[33] John S. Strong localization of photons in certain disordered dielectric superlattices. Phys. Rev. Lett., 1987, 58 (23): 2486.

[34] Winn J N, Fink Y, Fan S H, et al. Omnidirectional reflection from a one-dimensional photonic crystal. Opt. Lett., 1998, 23 (20): 1573.

[35] Felbacq D, Guizal B, Zolla F. Wave propagation in one-dimensional photonic crystals. Opt. Commu., 1998,152 (1—3): 119.

[36] Choi C G, Han Y T, Kim J T, et al. Air-suspended two-dimensional polymer photonic crystal slab waveguides fabricated by nanoimprint lithography. Appl. Phys. Lett., 2007, 90 (22): 1109.

[37] Combrie S, Weidner E, DeRossin A, et al. Detailed analysis by Fabry-Perot method of slab photonic crystal line-defect waveguides and cavities in aluminium-free material system. Opt. Exp. 2006,14 (16): 7353.

[38] Han S Z, Tian J, Feng S, et al. Fabrication of straight waveguide in two-dimensional photonic crystal slab and its light propagation characteristics. Acta Physica Sinica, 2005,54 (12): 5659.

[39] Subramania G, Lee Y J, Brener I, et al. Nano-lithographically fabricated titanium dioxide based visible frequency three dimensional gap photonic crystal. Opt. Exp., 2007,15 (20): 13049.

[40] Imada M, Lee L H, Okano M, et al. Development of three-dimensional photonic-crystal waveguides at optical-communication wavelengths. Appl. Phys. Lett., 2006,88 (17): 1107.

[41] Zeng Y, Y Fu, Chen X S, et al. Complete band gaps in three-dimensional quantum dot photonic crystals. Phys. Rev. B, 2006,74 (11): 5325.

[42] Kuchar F, Meisels R, Oberhumer P, et al. Microwave studies of photonic crystals. Adv. Eng. Mat., 2006,8 (11): 1156.

[43] Parimi P V, Lu W T, Vodo P, et al. Negative refraction and left-handed electromagnetism in microwave photonic crystals. Phys. Rev. Lett., 2004,92 (12): 7401.

[44] Vlasov Y A, O'Boyle M, Hamann H F, et al. Active control of slow light on a chip with photonic crystal waveguides. Nature, 2005,438 (7064): 65.

[45] Tanabe T, Notomi M, Mitsugi S, et al. All-optical switches on a silicon chip realized using photonic crystal nanocavities. Appl. Phys. Lett., 2005, 87 (15): 1112.

[46] L Zeng, Yi Y, C Hong, et al. Efficiency enhancement in Si solar cells by textured photonic crystal back reflector. Appl. Phys. Lett., 2006,89 (11): 1111.

[47] Mihi A, Lopez-Alcaraz F J, Miguez H. Full spectrum enhancement of the light harvesting efficiency of dye sensitized solar cells by including colloidal photonic crystal multilayers. Appl. Phys. Lett., 2006,88 (19): 3110.

[48] Skivesen N, Tetu A, Kristensen M, et al. Photonic-crystal waveguide biosensor. Opt. Exp., 2007,15 (6): 3169.

[49] Jensen J B, Pedersen L H, Hoiby P E, et al. Photonic crystal fiber based evanescent-wave sensor for detection of biomolecules in aqueous solutions. Opt. Lett., 2004, 29 (17): 1974.

[50] Joannopoulos J D, Johnson S G, Winn J N, et al. Photonic Crystals Molding the Flow of Light. 2nd edition. New Jersey: Princeton University Press, 2008.

[51] Johnson S G, Fan S H, Villeneuve P R, et al. Guided modes in photonic crystal slabs. Phys. Rev. B, 1999, 60(8): 5751.

[52] 何赛灵,戴道锌. 微纳光子集成. 北京: 科学出版社,2010.

[53] Johnson S G, Villeneuve P R, Fan S H, et al. Linear waveguides in photonic-crystal slabs. Phys. Rev. B, 2000, 62(12): 8212.

[54] Frandsen L H, Borel P I, Zhuang Y X, et al. Ultralow-loss 3-dB photonic crystal waveguide splitter. Opt. Lett., 2004, 29 (14): 1623.

[55] Settle D, Engelen P, Salib M, et al. Flatband slow light in photonic crystals featuring spatial pulse compression and terabertz bandwidth. Opt. Exp., 2007, 15(1): 219.

[56] Mori D, Baba T. Wideband and low dispersion slow light by chirped photonic crystal coupled waveguide. Opt. Exp., 2005, 13(23): 9398.

[57] Hou J, Gao D, Wu H, et al. Flat band slow light in symmetric line defect photonic crystal waveguides. IEEE Photo. Tech. Lett., 2009, 21: 1571.

[58] Hou J, Wu H, Citrin D S, et al. Wideband slow light in chirped slot photonic crystal coupled waveguides. Opt. Exp., 2010, 18(10): 567.

[59] Dai L, Wang F, Jiang C. Flatband slow wave in novel air-hole-array strip waveguides. IEEE Photo. J., 2009, 1(3): 178.

[60] Liu C Y, Chen L W. Tunable photonic-crystal waveguide Mach - Zehnder interferometer achieved by nematic liquid-crystal phase modulation. Opt. Exp., 2004, 12(12): 2616.

[61] Jiang Y Q, Jiang W, Gu L L, et al. 80-micron interaction length silicon photonic crystal waveguide modulator. Appl. Phys. Lett., 2005,87(22): 1105.

[62] Zucker J E, Jones K L, Chiu T H, et al. Strained quantum wells for polarization-independent electrooptic waveguide switches. J. Lightwave Technology, 1992,10 (12): 1926.

[63] Jenkins R M, Devereux R W J, and Heaton J M. A novel waveguide Mach-Zehnder interferometer based on multimode interference phenomena. Opt. Commu., 1994,110: 410.

[64] Lagali N S, Paiam M R, et al. Analysis of generalized Mach-Zehnder interferometers for variable-ratio power splitting and optimized switching. J. Lightwave Technology, 1999, 17(12): 2542.

[65] Ulrich R and Ankele G. Self-imaging in homogeneous planar optical waveguide. Appl. Phys. Lett., 1975, 27(6), 337.

[66] Soldano L B, Pennings E C M. Optimal multimode interference devices based on self-imaging: principles and applications. J. Lightwave Technology, 1995, 13 (4): 615.

[67] Dai D, He S. Optimization of ultracompact polarization-insensitive multimode interference couplers based on Si nanowire waveguides. IEEE Photo. Tech. Lett., 2006, 18(19): 2017.

[68] Tsao S L, Guo H C, Tsai C W. A novel 1×2 single-mode 1300/1550 nm wavelength division multiplexer with output facet-tilted MMI waveguide. Opt. Commu., 2004, 232: 371.

[69] Smith M K. New focusing and depressive planar component based on an optical phased array. Electr. Lett., 1988, 24(7), 385.

[70] Vellekoop A R, Smith M K. Four-channel integrated optic wavelength demultiplexer with weak polarization dependence. J. Lightwave Technology, 1991, 9 (3): 310.

[71] Takahashi H, Suzuki S, Kato K, et. al. Arrayed-waveguide grating for wavelength division multi/demultiplexer with nanometer resolution. Electr. Lett., 1990, 26(2): 87.

[72] Dragone C. An N × N optical multiplexer using a planar arrangement of two star couplers. IEEE Photo. Tech. Lett., 1991, 3(9): 812.

[73] Smit M K, Dam C V. Phasar-based WDM-devices: principles, design and applications. IEEE J. Quant. Electr., 1996, 2(2): 236.

[74] Doi Y, Ishii M, Kamei S, et al. Flat and high responsivity CWDM photoreceiver using silica-based AWG with multimode output waveguides. Electr. Lett., 2003, 39(22): 1603.

[75] Bogaerts W, Taillaert D, Dumon P, et al. A polarization-diversity wavelength duplexer circuit in silicon-on-insulator photonic wires. Opt. Exp., 2007, 15(4): 1567.

[76] Marcatili E A J. Bends in optical dielectric guides. Bell Syst. Tech. J., 1969, 48(9): 2103.

[77] Oda K, Suzuki S, Takahashi H, et al. An optical FDM distribution experiment using a high finesse waveguide-type double ring resonator. IEEE Photo. Lett., 1994, 6(8): 1031.

[78] Yi H, Citrin D S, Zhou Z. Highly sensitive silicon microring sensor with sharp asymmetrical resonance. Opt. Exp., 2010,18: 2967.

[79] Heebner E, Chak P, Pereira S, et al. Distributed and localized feedback in microresonator sequences for linear and nonlinear optics. J. Opt. Soc. Am. B, 2004, 21(10): 1818.

[80] Absil P, Hryniewicz V, Little B E, et al. Compact microring notch filters. IEEE Photo. Tech. Lett., 2000, 12(4): 398.

[81] Yan X, Ma C S, Xu Y Z, et al. Analysis and optimization for a polymer cross-grid array of microring resonant wavelength multiplexers. Opt. Eng., 2005, 44(7): 075001.

[82] Almeida V R, Barrios C A, Panepucci R R, et al. All-optical control of light on a silicon chip. Nature, 2004, 431(10): 1081.

[83] Xu Q F, Schmidt B, Pradhan S, et al. Micrometre-scale silicon electro-optic modulator. Nature, 2005, 435(5): 325.

[84] Madsen C K, Lenz G. Optical all-pass filters for phase design with application for dispersion compensation. IEEE Photo. Tech. Lett., 1998, 10(7): 994.

[85] Absil P P, Hryniewicz J V, Little B E, et al. Wavelength conversion in GaAs microring resonators. Opt. Lett., 2000, 25(8): 554.

[86] White I M, Oveys H, Fan X D. Liquid-core optical ring-resonator sensors. Opt. Lett., 2006, 31(9): 1319.

[87] Chao C Y, Fung W, Guo L J. Polymer microring resonators for biochemical sensing applications. IEEE J. Sel. Topic. Quant. Electr., 2006, 12(1): 134.

[88] 夏志轩. 光学微环生物传感器的设计与优化[硕士学位论文]. 武汉: 华中科技大学, 2008.

[89] Yi H, Citrin D S, Chen Y, et al. Dual-microring-resonator interference sensor. Appl. Phy. Lett., 2009, 95(19): 1112.

[90] Turner A C, Foster M A, Gaeta A L, et al. Ultra-low power parametric frequency conversion in a silicon microring resonator. Opt. Exp., 2008, 16(7): 4881.

[91] Lu Y, Liu F, Qiu M. All-optical format conversions from NRZ to BPSK and QPSK based on nonlinear responses in silicon microring resonators. Opt. Exp., 2007, 15(21): 14275.

[92] Van V, Ibrahim T A, Absil P, et al. Optical signal processing using nonlinear semiconductor microring resonators. IEEE J. Sel. Topic. Quant. Electr., 2002, 8: 705.

第五章　硅基光源

硅基光源是硅基光电子学元器件中的重中之重.由于硅是间接带隙的半导体,发光效率不高,因此硅一直以来被认为不适合制作光源材料,但由于其巨大的潜在价值,人们一直没有放弃在硅上制备激光器的努力,从20世纪90年代发现多孔硅的室温发光,2000年观察到纳米硅的增益,直到最近几年硅拉曼激光器、Ⅲ-Ⅴ族-硅混合激光器和锗硅激光器,无不引起世界光电子研究学者的极大关注,给大家显示出成功制备硅基激光器的曙光.本章将首先介绍光发射的普适理论,解释块体硅不能制备激光器的原因.然后,从材料角度出发,介绍三种提高硅发光效率的方法：量子效应低维硅、稀土掺杂、利用能带理论设计直接带隙材料.在上述材料的基础上介绍了三种发光二极管：体硅发光二极管、纳米硅发光二极管、稀土掺杂硅和稀土硅酸盐化合物发光二极管.最后介绍了最新硅基激光器的进展,包括硅拉曼激光器、Ⅲ-Ⅴ族-硅混合激光器和锗硅激光器.从而能够帮助读者理解硅基发光的基本原理以及当前硅基光源的研究重点及方向,探明硅基光电子系统中光源的研究现状及未来的发展趋势.

§5.1　光发射的基础理论

5.1.1　光辐射的理论[1~3]

1. 辐射复合

根据能量守恒原则,电子和空穴复合时应释放一定的能量,如果能量是以光子的形式放出的,这种复合称为辐射复合.如果能量以热能、动能等形式放出,则称为非辐射复合.辐射复合又分为带间复合和非带间复合.带间复合是指导带电子与价带的空穴直接复合;非带间复合是通过复合中心进行,包括有杂质或缺陷参与,如电子从导带跃迁到杂质能级,或杂质能级上的电子跃迁入价带,或电子在杂质能级之间的跃迁,其中带间复合是辐射复合中的主要形式.

对于直接带隙半导体,导带和价带的极值都在 $\boldsymbol{k}$ 空间原点,带间复合为直接跃迁,由于直接跃迁的发光过程只涉及一个电子-空穴对和一个光子,其辐射效率很高.对于间接带隙半导体,导带和价带的极值对应于不同的波矢 $\boldsymbol{k}$,这时发生的带与带之间的跃迁是间接跃迁.在间接跃迁过程中,除了出射光子外,还需要声子参与,因此间接跃迁比直接跃迁的概率小得多.

2. 非辐射复合

非辐射复合的本质就是将电子和空穴复合释放的能量以非光子的形式释放，如转变为热能或其他形式的能量，其中主要有俄歇复合和多声子复合两种方式.

(1) 俄歇复合. 一个电子或空穴吸收另一电子与空穴复合时放出的能量，跃迁到更高的能量状态的过程称为俄歇复合过程. 在俄歇复合过程中获得能量而跃迁到更高能态的载流子与周围晶格反复碰撞而失去能量的几率很大. 这种过程的几率与复合的载流子浓度和接受能量的载流子浓度乘积成正比，所以载流子浓度高的材料俄歇复合过程更容易. 因此，发光器件的掺杂浓度不能太高，如果太高，俄歇复合过程比较严重.

(2) 多声子复合. 晶体中电子和空穴复合时可以发射多个声子来释放能量称之为多声子复合. 通常发光材料的带隙较宽，均在 1 eV 以上，而一个声子的能量大约为 0.06 eV，所以同时发射多个声子的几率是很小的. 但是晶体中存在许多杂质和缺陷，在带隙中分布着许多分立的能级，电子依次落入这些能级和连续发射多个声子仍然是可能的，最终可以释放大的能量.

3. 自发辐射

辐射复合又分为自发辐射和受激辐射. 自发辐射是指处于激发态的原子中的电子在激发态能级上只能停留一段很短的时间，就自发地跃迁到较低能级中去，同时辐射出一个光子.

这一概念最早是由爱因斯坦 1916 年提出. 如图 5.1 所示：E_1 是基态，E_2 是激发态. 电子在这些能级的跃迁必定伴随着吸收或发射频率为 $h\nu$ 的光子. 在常温下，大部分原子都处于基态，如有能量为 $h\nu=E_2-E_1$ 的光子与原子系统相互

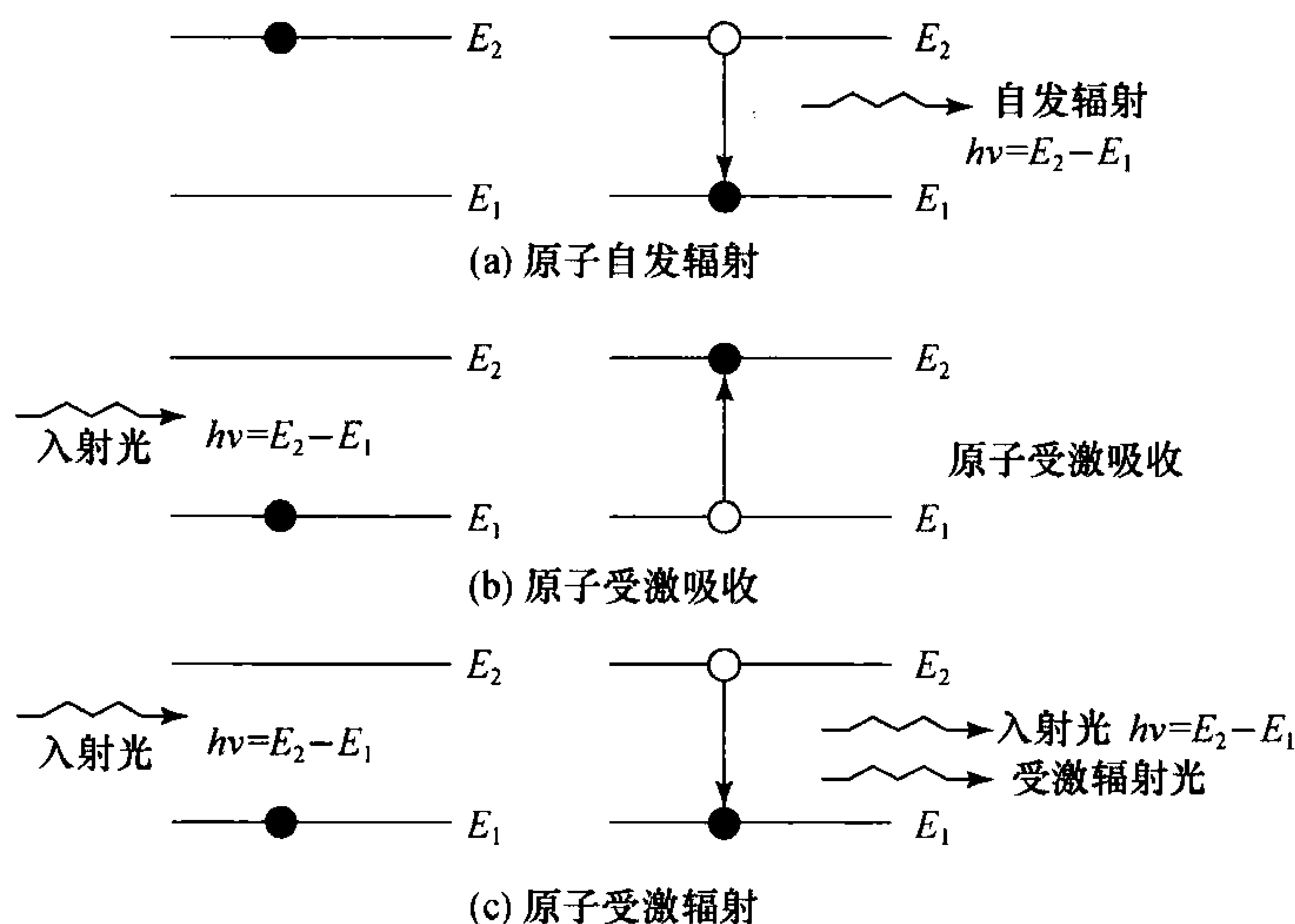


图 5.1 原子辐射跃迁示意图

作用,则处于基态的原子吸收光子进入激发态 E_2. 激发态是不稳定的,经过短时间后,原子必然跃迁回到基态 E_1,同时发射能量为 $h\nu$ 的光子,这种不受外界因素的作用,原子自发的从激发态回到基态引起的光子发射过程,称为自发辐射. 半导体的自发辐射寿命变化很大,典型值为 $10^{-9}\sim10^{-3}$ s,它取决于各种半导体参数,如禁带宽度及复合中心浓度等.

4. 受激辐射

当原子处于激发态 E_2 时,如果恰好有能量等于 E_2-E_1 的光子入射进来,在入射光子的影响下,原子会发出一个同样的光子而跃迁到低能级 E_1 上去,这种辐射叫做受激辐射. 这种方式就是产生激光的基本原理.

自发辐射和受激辐射是两种不同的光子发射过程,自发辐射中各原子的跃迁都是随机的,所产生的光子虽然具有相等的能量 $h\nu$,但这种光辐射的位相和传播方向等都各不相同. 受激辐射却不同,它所发出的光子的全部特性(频率、位相、方向和偏振态等)同入射光子完全相同. 另外,自发辐射过程中,原子从 E_2 态跃迁到 E_1 态,伴随着发射一个光子,而在受激辐射过程中,一个入射光子 $h\nu$ 使激发态从 E_2 态跃迁到 E_1 态,同时发射出两个同相位、同频率的 $h\nu$ 光子. 受激辐射光的频率、相位、偏振态和传播方向与入射光相同,这种光称为相干光. 自发辐射光是由大量不同激发态的电子自发跃迁产生的,其频率和方向分布在一定范围内,相位和偏振态是混乱的,这种光称为非相干光.

5. 受激吸收

在正常状态下,电子处于基态 E_1,在入射光作用下,它会吸收光子的能量跃迁到激发态 E_2 上,这种跃迁称为受激吸收. 电子跃迁后,在基态留下相同数目的空穴. 受激吸收是受激辐射的逆过程. 设在单位物质中,处于 E_1 和 E_2 的原子数分别为 N_1 和 N_2. 当系统处于热平衡状态时,存在下面的分布

$$\frac{N_2}{N_1}=\exp\left(-\frac{E_2-E_1}{kT}\right), \tag{5.1.1}$$

式中 $k=1.381\times10^{-23}$ J/K,为玻尔兹曼常数,T 为热力学温度. 由于 $E_2-E_1>0$, $T>0$,所以在这种状态下,总是 $N_1>N_2$. 这是因为电子总是首先占据低能量的轨道. 如果 $N_1>N_2$,即受激吸收大于受激辐射. 当光通过这种物质时,光强按指数衰减,这种物质称为吸收物质. 如果 $N_2>N_1$,即受激辐射大于受激吸收,当光通过这种物质时,会产生放大作用,这种物质称为激活物质. $N_2>N_1$ 的分布和正常状态($N_1>N_2$)的分布相反,所以称为粒子数反转.

6. 发光效率

因为发光过程中同时存在辐射复合和非辐射复合过程. 两者复合概率的不同使材料具有不同的发光效率. 发射光子的效率决定于载流子辐射复合寿命 τ_r 和非辐射复合寿命 τ_{nr} 的相对大小. 通常用"内量子效率"$\eta_内$ 和"外量子效率"$\eta_外$

来表示发光效率."内量子效率"为单位时间内辐射复合产生的光子数与单位时间内注入的电子-空穴对数之比,也就是辐射复合速率 w_r 与总的复合速率 $w_{nr}+w_r$ 之比.而辐射复合速率和非辐射复合速率又分别为辐射复合寿命 τ_r 和非辐射复合寿命 τ_{nr} 的倒数,即

$$\eta_{内} = \frac{w_r}{(w_{nr}+w_r)} = \frac{\tau_{nr}}{(\tau_{nr}+\tau_r)} = 1-\frac{\tau_r}{(\tau_{nr}+\tau_r)}, \tag{5.1.2}$$

其中

$$w_t = w_r + w_{nr} = \frac{1}{\tau_r}+\frac{1}{\tau_{nr}}.$$

因此,只有当 $\tau_{nr} \gg \tau_r$ 时,才能获得高效率的光子发射.

对间接复合为主的半导体材料,一般既存在发光中心,又存在其他复合中心,通过前者产生辐射复合,后者产生非辐射复合.因此,要使辐射复合占优势,必须使发光中心浓度远大于其他杂质浓度.

外量子效率是指半导体材料中总的有效发光效率.单位时间内发射到晶体外部的光子数与单位时间内注入的电子-空穴对数之比.因为很多辐射复合所产生的光子并不是全部都能离开晶体向外发射.从发光区产生的光子通过半导体时部分可以被吸收,另外由于半导体的高折射率,光子在界面处很容易发生全反射而返回晶体内部.因此一般外量子效率比内量子效率低很多.

5.1.2 光放大和增益

光的放大主要由材料的增益谱决定,对于半导体材料,它是由态密度 $\rho(h\nu)$、费米转换因子 $f_g(h\nu)$ 和辐射寿命 τ_r 决定的[4].

$$\begin{aligned}\mathrm{d}\Phi(h\nu) &= \mathrm{d}r_{stim}(h\nu) - \mathrm{d}r_{abs}(h\nu) \\ &= \frac{\lambda^2}{8\pi\tau_r}\rho(h\nu)f_g(h\nu)\Phi(h\nu)\mathrm{d}z \\ &= g(h\nu)\Phi(h\nu)\mathrm{d}z,\end{aligned} \tag{5.1.3}$$

其中 $\mathrm{d}r_{stim}$ 和 $\mathrm{d}r_{abs}$ 是一定光子能量 $h\nu$ 下的受激发射和受激吸收率,Φ 是光子流密度,$g(h\nu)$ 是增益系数,$\mathrm{d}\Phi$ 是光子流量的变化,且

$$f_g(h\nu,E_F^e,E_F^h,T) = f_e(h\nu,E_F^e,T) - [1-f_h(h\nu,E_F^h,T)], \tag{5.1.4}$$

其中 f_e 和 f_h 是电子-空穴对的热分布函数,E_F^e 和 E_F^h 是电子和空穴的准费米能级,当没有外泵浦的情况下,费米转换因子减少到简单的费米态,也就是对于一个空的导带和填满的价带,增益系数小于吸收系数,$f_g<0$.当外泵浦激发高密度的自由载流子时,准费米能级的劈裂增加,当 $E_F^e-E_F^h>h\nu$ 时,满足粒子数反转条件,$f_g>0$.这意味上面的公式为正值,因此系统也显示正的增益.从上面公式可以看出,辐射寿命 τ_r 也是一个关键的参数,寿命越短,增益越大.

对于一个原子系统，增益系数的表达式简化为

$$g(h\nu) = \sigma_{em}(h\nu)N_2 - \sigma_{abs}(h\nu)N_1, \tag{5.1.5}$$

其中 σ_{em}是发射截面，σ_{abs}是吸收截面，N_2 和 N_1 分别代表激发态和基态能级中有源物质的粒子数. 如果发射截面和吸收截面相等，有正向增益的条件是 $N_2 > N_1$，也就是实现粒子数反转.

同样，如果光在有源材料中传输时，光强会因光吸收而呈指数关系下降，设 α_p 为光吸收系数，只要这个材料的增益系数大于吸收系数，$g > \alpha_p$，也就是增益大于损耗，就可以获得光放大. 如果该系统是长度为 L 波导，I_T 和 I_o 为出射和入射的光强度，则光放大因子由下式表达：

$$G = \frac{I_T}{I_o} = \exp[(\Gamma g - \alpha_p)L] > 1, \tag{5.1.6}$$

其中 Γ 有源物质区光模场的限制因子.

对于激光器来讲，如果光反馈由法布里-珀罗(Fabry-Perot, F-P)腔提供，当通过 F-P 腔往返增益(整体增益经历了由一个横渡来回光子腔)大于 1. 这个条件可以表示为关系

$$G^2 R_1 R_2 > 1,$$

其中 R_1 和 R_2 分别为背部和前面的反射率.

5.1.3 激光器的三大要素[1,3]

1. 激光工作物质

用来实现粒子数反转并产生光的受激辐射放大作用的工作物质，称为激光工作物质. 对激光工作物质的主要要求，是尽可能在其工作粒子的特定能级间实现较大程度的粒子数反转，并使这种反转在整个激光发射作用过程中尽可能有效地保持下去，为此，要求工作物质具有合适的能级结构和跃迁特性. 对于半导体来说，直接带隙半导体的发光效率比间接带隙半导体高 3 个数量级，因此直接带隙半导体材料更适合制作激光器.

2. 粒子数反转

在通常的情况下，高能态中的粒子数目总是比低能态中的粒子数目少许多. 通过光照、高压放电、电流注入、化学反应等不同方式，将光能、电能、化学能等不同方式的能量转给激光物质，使激光物质中的粒子由低能态抽运到高能态. 实现了能量转换. 如果泵浦的能力足够强，使位于高能态的粒子数目远远大于该温度下平衡态时低能态的粒子数，从而实现粒子数的反转. 假定有两个能级 E_1 和 E_2，并且 $E_2 > E_1$，它们上面的电子数目分别为 N_1 和 N_2，正常情况下 $N_1 > N_2$，如果采用电注入或其他泵浦方式使得 $N_2 > N_1$，则实现了粒子数反转. 根据工作物质和激光器运转条件的不同，可以采取不同的激励方式和激励装置，常见的有

以下四种：① 光学激励(光泵)，即利用外界光源发出的光来辐照工作物质以实现粒子数反转；② 气体放电激励，即利用在气体工作物质内发生的气体放电过程来实现粒子数反转；③ 化学激励，即利用在工作物质内部发生的化学反应过程来实现粒子数反转，通常要求有适当的化学反应物和相应的引发措施；④ 核能激励，即利用小型核裂变反应所产生的裂变碎片、高能粒子或放射线来激励工作物质并实现粒子数反转.

3. 谐振腔

谐振腔对频率一定、方向一致的光产生正反馈，使其获得足够大的增益，克服内部和端面的损耗，从而发生谐振，产生激光振荡. 因此谐振腔起到了产生反馈、选择波长、形成谐振、发出激光的作用. 其通常是由具有一定几何形状和光学反射特性的两块反射镜按特定的方式组合而成. 作用是提供光学反馈能力，使受激辐射光子在腔内多次往返以形成相干的持续振荡. 其通常由组成腔的两个反射镜的几何形状(反射面曲率半径)和相对组合方式所决定；对腔内往返振荡光束的方向和频率进行限制，以保证输出激光具有一定的定向性和单色性. 则是由给定共振腔型对腔内不同行进方向和不同频率的光，具有不同的选择性损耗特性所决定.

§5.2　硅放大的限制

5.2.1　硅的间接带隙

目前硅的激光器还一直没有实现，仍然是世界难题. 为什么硅不适合制备激光器. 下图给出了硅的能带结构[4]. 由图 5.2 可见，硅导带和价带的极值对应于

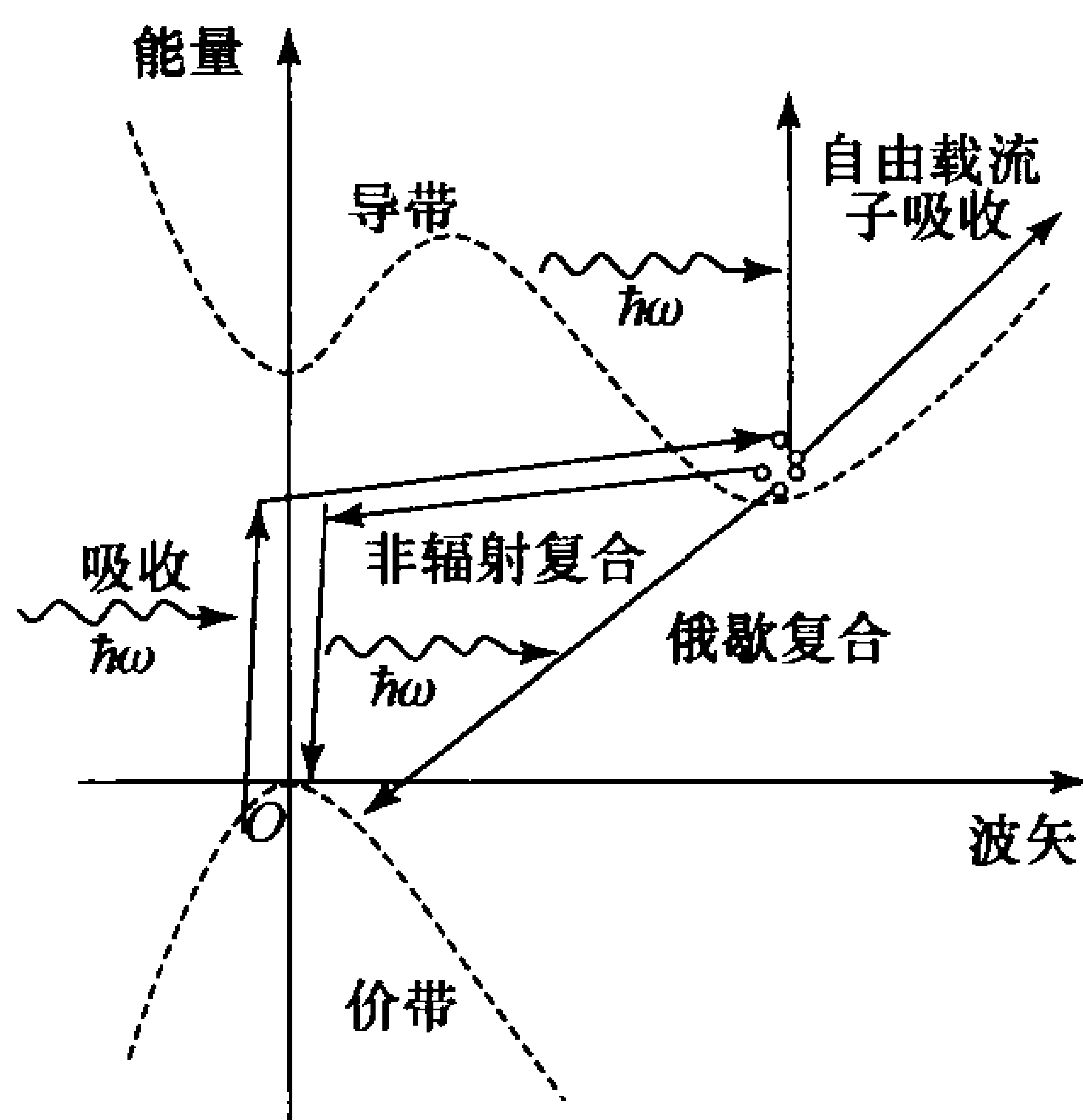


图 5.2　硅的能带结构

不同的波矢 k，是间接带隙的半导体，辐射复合几率很低，同时存在两个强非辐射跃迁过程：俄歇复合和自由载流子吸收. 电子-空穴对的辐射寿命长(毫秒量级)，一个电子-空穴对需要毫秒才能复合. 在此期间，电子和空穴移动的体积达到 10 μm^3. 这样他们很容易遇到缺陷或俘获中心，载流子就会发生非辐射复合. 在硅中，典型的非辐射寿命是纳秒量级，因此内量子效率 η 约为 $10^{-6}\sim10^{-5}$. 这是为什么硅是一个弱的发光材料的原因，也就是有效的非辐射复合快速的耗尽激发的载流子. 许多战略性的研究正在开展来克服这种硅限制，如杂质掺杂、量子限制、硅-锗合金、超点阵等方案.

5.2.2 硅的俄歇复合和自由载流子吸收

除了硅的间接带隙限制外，还有两个现象限制了硅在光学放大中的应用.

第一个是无辐射三粒子复合机理，一个激发的电子-空穴复合将多余的能量释放给另一个电子(空穴)，这就是所谓的非辐射俄歇复合机制. 一旦多的载流子被激发，这种机理就非常严重. 一个俄歇复合的概率是和激发的载流子数 Δn 的平方成正比，和禁带宽度成反比. 因为半导体中有很多的载流子，所以俄歇复合是很强的. 俄歇复合造成的非辐射复合寿命为[5]

$$\tau_A = 1/C\Delta n^2, \tag{5.2.1}$$

其中 C 为常数，和材料的掺杂浓度有关. 对于硅，$C=10^{-30}\ cm^6\cdot s^{-1}$，当 $\Delta n=10^{19}\ cm^{-3}$时，非辐射复合寿命为 10 ns[6]. 因此，对于高载流子注入硅，俄歇复合是非辐射复合的主要机制.

第二种限制是自由载流子吸收，激发的载流子可以吸收光子，因此消耗了反转的粒子数. 同时增加了由于信号引起的光学损耗. 自由载流子吸收系数是和硅的自由载流子浓度 n_{fc}以及光波长 λ 有关. 在 300 K 时，$\alpha_n\sim10^{-18}n_{fc}\lambda^2$. 若 $n_{fc}=10^{19}\ cm^{-3}$，$\lambda=1.55\ \mu m$，则 $\alpha_n=24\ cm^{-1}$.[6]对于重掺杂硅，这也是产生激光的主要限制，然而对于本征硅，除非 n_{fc}非常高，否则这种贡献比较小.

§5.3 硅基发光材料

近二十年，人们利用多个途径来提高硅基材料的发光效率. 比如说，利用低维硅的量子效应，包括多孔硅、纳米硅、纳米结构硅；利用杂质掺杂，如铒掺杂硅；利用能带工程变成直接带隙，如锗、硅材料；等等.

5.3.1 低维硅[7]

1990 年，Canham 等[8]报道了多孔硅(实际是一种纳米材料)在室温下可以辐射较强的可见光. 从此，以深入研究多孔硅发光性能为起点，人们开始了对硅基

光电子学光源的研究热潮.最早的制备多孔硅的方法是采用电解液电化学腐蚀氧化硅,形成含有孔洞和纳米硅晶粒的表面,硅表面形成的多孔状结构称为多孔硅.电化学腐蚀形成多孔硅的孔径一般为10～50 nm,残留硅颗粒直径为2～8 nm,电子衍射图像表明,多孔硅晶粒保持了一定的有序性,晶向与单晶硅基底一致.多孔硅辐射的光波长在可见光范围,对应的光子能量比体硅禁带宽度高,且随着多孔硅晶粒尺寸减小,光波长逐渐蓝移.多孔硅发光强度比体硅高3～4个数量级.

目前对低维纳米硅基材料发光机理的研究出现了各种各样的物理模型,各有优缺点,总体上看,目前主要存在以下3种发光机理模型:

① 量子限制发光模型.量子理论认为纳米半导体材料中的载流子将受到晶粒势垒的限制,称为量子限制效应.当硅晶粒尺寸逐渐减少到几个纳米时,由于周期性晶体中平移对称性的丧失,使得间接带隙对于光的选择定制不再适用.就能带结构而言,晶粒尺寸小到几个纳米的硅晶粒,其禁带宽度将会明显增加.尤其是当晶粒尺寸 $d<3$ nm 时,其禁带宽度与 d^{-2} 近似呈线性依赖关系[9].与此同时,由于能级的分裂将在导带和价带中出现量子化的能级.而且晶粒尺寸越小,量子化能级的分裂会越加显著.硅能带结构的这一实质性变化,将会对其发光特性产生重要影响,即随着硅晶粒尺寸的减小,不仅使其发光强度增加,而且还出现谱峰蓝移现象[10].量子限制效应发光的本质是被封闭在纳米量子点中的载流子,在光泵浦或电注入条件下,从价带被激发到导带较高的量子化能级上,而后再从这些能级热弛豫到较低的量子化能级上,并与价带中的空穴直接发生辐射复合而导致光发射的.其发光波长直接与纳米晶粒的带隙能量和量子化能级间距密切相关.

② 与氧有关的缺陷发光模型.存在于半导体中的缺陷或杂质,会在禁带中引入相应的缺陷与杂质能级,并由此形成发光中心,当电子被光能或电能激发后会落入该发光中心,而后,当该电子跃迁到价带能级时便会发光.目前所制备的各种硅基纳米发光材料,多系由 SiO_2、含 H 的 SiO_x 以及 α-Si:H 包围的镶嵌结构.因此膜层中除了硅晶粒外,其中的氧原子将与硅原子或氢原子以各种价键形式相结合,并产生各类与氧相关的缺陷.这些新生缺陷将在 SiO_2 或 SiO_x 禁带中引入相应的局域能级,从而对其发光特性产生重要的影响[11,12].

③ 量子限制效应-发光中心发光模型[13].前面,我们简要介绍了量子限制效应和与氧相关的缺陷发光特性.然而在很多情况下,硅基纳米结构的光致发光(Photoluminescence, PL)或电致发光(electroluminescence, EL)特性,并非单由其中的某一种发光机制所支配,而是有两者共同决定.尤其是对于由 SiO_2 膜覆盖的Si纳米晶粒或Si量子点,经表面氧化处理的纳米多孔硅,由薄 α-Si 层与 SiO_2 层形成的超晶格或量子阱结构更是如此.但应该指出的是,这种结构与氧相关的缺陷不是非桥键氧缺陷中心或氧空位等,而是位于 Si/SiO_2 界面区域内

的局域表面态.他们对于 Si 基纳米材料发光特性的贡献,不是直接起着一个发光中心的作用,而是扮演着一种辐射复合中心的角色.

一般而言,Si 氧化物中的表面局域态或缺陷的形成可以影响 Si 基纳米结构 PL 峰值能量和形状,而不影响其红移或蓝移特性.正如人们所熟知的那样,这些局域态是由界面原子的无序性(如结构无序或组分无序等)所引起,它们自身位于纳米晶粒和包围这些晶粒的介质层之间,而能量是处于纳米晶粒的禁带之中.对于由不同物质覆盖钝化的表面,其局域能量也不相同.如当纳米晶粒表面由 H 原子进行表面钝化时,对于一个粒径尺寸为 1.67 nm 的纳米晶粒而言,其表面晶格弛豫能量为 86 meV.然而,当电子-空穴对被局域在表面二聚物的一个具有弱相互作用的 Si 悬键上时,其局域能量可高达 0.52 eV.更令人感兴趣的是,在 Si/SiO_2 界面形成的二聚物提供了一个比由 H 终端的表面更强的局域化效应.因此,对于纳米晶粒而言,表面无序和畸变的诱导是产生表面态的直接原因.按照纳米材料的量子化效应,随着晶粒尺寸的减小,其表面-体积比会迅速增加,因此对于小晶粒尺寸的纳米结构而言,它将显著影响其 PL 特性.

5.3.2　铒掺杂硅和铒硅酸盐化合物

在硅中掺入一定浓度的稀土元素,充分利用稀土元素丰富的能级结构发出不同波长的光来满足实际的需要.其中稀土铒可以发出 1.53 μm 的光,对应于光纤通信中的石英玻璃吸收的最小值,且该波长的能量不受激发功率和所处环境温度的影响,是硅基光电子学的标准波长.因而,掺 Er 硅发光是一种很有发展前途的硅基发光材料.目前有两个方向正在开展,一个是掺 Er 富硅氧化硅(氮化硅)[14~16],另一个是用 Er 的硅酸盐化合物[17~22].

富硅氧化硅中的硅纳米晶是一个很好的 Er 敏化剂,能量转移效率可以达到 70%.目前报道的最好结果是韩国 Shin 教授研究小组[23]制备掺 Er 富硅氧化硅放大器,在 1 cm 长的波导中获得了 4 dB 的光增益.但该系统还存在一些问题,如 Er 和硅纳米晶的耦合仍然不完全,报道仅有 5%的 Er 被耦合纳米硅晶,其他通过 Er 的直接吸收激发.而且,光增益和纳米硅晶的浓度密切相关,控制纳米硅晶尺寸和浓度的工艺较复杂;并且由于纳米硅晶的存在,它受载流子吸收的限制,即使通过精确控制 Er 和纳米硅晶的浓度,吸收仍很大,很难获得高增益.另外其采用的泵浦方式是 477 nm 的氩(Ar)离子激光器光泵浦的,离我们需要的电泵浦的要求还有一定距离.

在 Er 的硅酸盐化合物结构中,Er 是化合物的阳离子,而不再是作为杂质掺杂进去,和以前掺杂方法相比,成功将 Er^{3+} 浓度提高了 2 个数量级,达到 $10^{22}\ cm^{-3}$,而且由于 Er^{3+} 是固溶在化合物中的,减少了 Er^{3+} 之间的团聚现象. Isshiki 等[17]采用溶胶-凝胶法制备了一种 Er 的硅酸盐化合物.图 5.3(a)为这种

材料的高分辨透射电子显微镜(transmission electron microscope, TEM)图和光致发光谱. 由图可见，这种材料具有较少的缺陷，并有强的 1.528 μm 室温光致发光. 但这种材料由于具有非常高的 Er 离子浓度，合作上转换明显，所以采用钇(Y)和镱(Yb)共溶的方法分散 Er，因为 Y(Yb)的离子半径和 Er 接近，所以可以使 Er 充分分散. 结果显示当 Er ∶ Y 比例为 1 ∶ 19，光致发光强度提高了 30 倍[21]. 另外，采用 4 能级模型，通过 Er 速率方程和传输方程数值计算了光波导 Er-Y 硅酸盐化合物放大器光增益与 Er^{3+} 浓度、波导长度、泵浦功率和信号功率的关系，如图 5.3(b)所示. Er^{3+} 浓度为 0.5×10^{22} cm^{-3}，当泵浦功率为 70 mW 时，在 1 mm 波导长度下可以获得 10 dB 以上的增益[24]. 在理论计算的基础上，制备的 7.8 mm 长的脊形波导上观察到了 5.5 dB 的光放大[25]. 但由于 Er 的硅酸盐化合物的形成机理比较复杂，对工艺参数的依赖性很大，微小的氧分压、基体和温度等参数的差别都会导致不用相结构的形成，Er 的硅酸盐化合物光源器件还需要不断的优化.

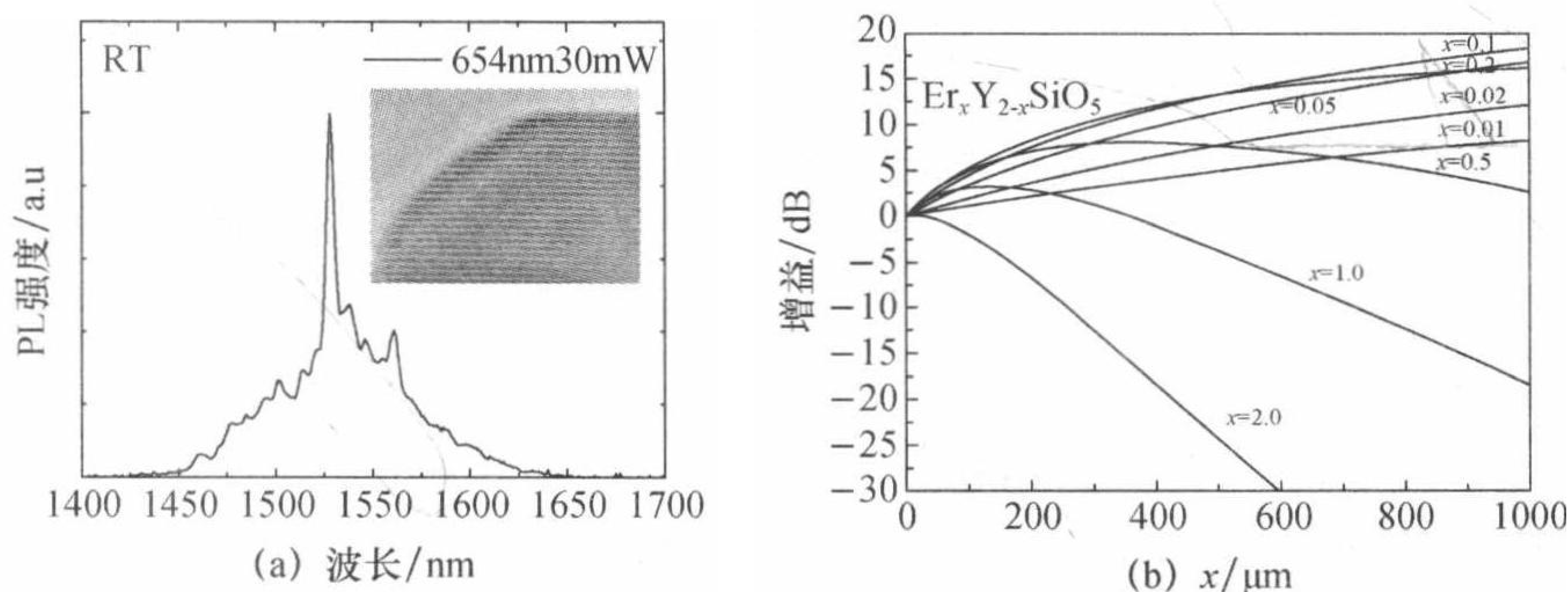


图 5.3 Er 的硅酸盐化合物材料的高分辨 TEM 图和 PL 谱(a)；Er 的硅酸盐化合物材料的光增益(b)

掺铒硅发光效率的增强一直是人们关注的热点，最近几年人们研究出了几种新的结构和方法来提高掺铒硅发光效率，如表面等离子辅助增强铒发光[26,27]、掺铒环形微腔激光器[28,29]、光子晶体增强铒发光[30]、槽形波导(slot waveguide)宽带发光增强效应[31]等.

表面等离子体的辅助作用可以有效改变 Er 离子的发光光谱，采用二维金属光子晶体可以提高 Er^{3+} 的发光效率[27]. 由于金属光子晶体可以被局限在很小的范围内，例如在金属光子晶体的能带边缘处，表面等离子体的群速率接近零，因此相应的态密度较其他位置可提升数个量级. 通过将处于能带边缘处的表面等离子体耦合为光子并发射到远场，Er^{3+} 在特征波长的发光效率可以显著增强，Wang 等[27]设计了一个掺铒发光器件模型，发现二维金属光子晶体能带边缘的表面等离子体的态密度较平面金属增强了 40 倍，发光效率也提高近 40 倍.

Polman 等[28]采用标准硅工艺结合高 Q 腔来提高 Er 的发光效率，并且制备出掺 Er 环形微腔激光器. 环形微腔直径为 40 μm，当泵浦波长为 1460 nm 时，出射的激光波长为 1578 nm，阈值功率为 4.5 μW，相应 Q 值为 3.9×10^7. 2009 年，他们又在上面实验的基础上考虑 Er 的上转换，报道了绿光激光器[29]. 最近，Makarova 等[30]采用光子晶体纳米腔增强 Er 的发光效率. 他们首先在 SOI 基体沉积 100 nm 的掺铒富硅氮化硅，然后采用电子束曝光刻蚀掺铒富硅氮化硅以及 SOI 的硅晶体，最后 SOI 中的氧化层用氢氟(HF)酸去除，形成光子晶体纳米腔和掺铒富硅氮化硅的复合结构，用 473 nm 激光器激发，通过改变光子晶体的孔径比可以提高 PL 强度达到 20 倍以上，Q 值为 6000.

最近，槽形波导引起人们的关注，它包括一个纳米尺度的低折射率区域和两侧高折射率区域. 由于电位移的连续性，低折射率区域可以获得高的电场强度. 在高电场强度下可以提高低折射率区域的光发射效率. 斯坦福大学 Brongersma 小组[31]研究了硅的槽形波导增强发光. 图 5.4 为一个槽形波导的示意图，当中间的低折射率层为 20 nm，宽度为 300 nm，高为 410 nm 时，y 方向的自发辐射可以提高 20～30 倍，当改变低折射率层到 30 和 50 nm 时，放大倍数降低. 这种新的结构为我们开辟了一个提高发光的新途径.

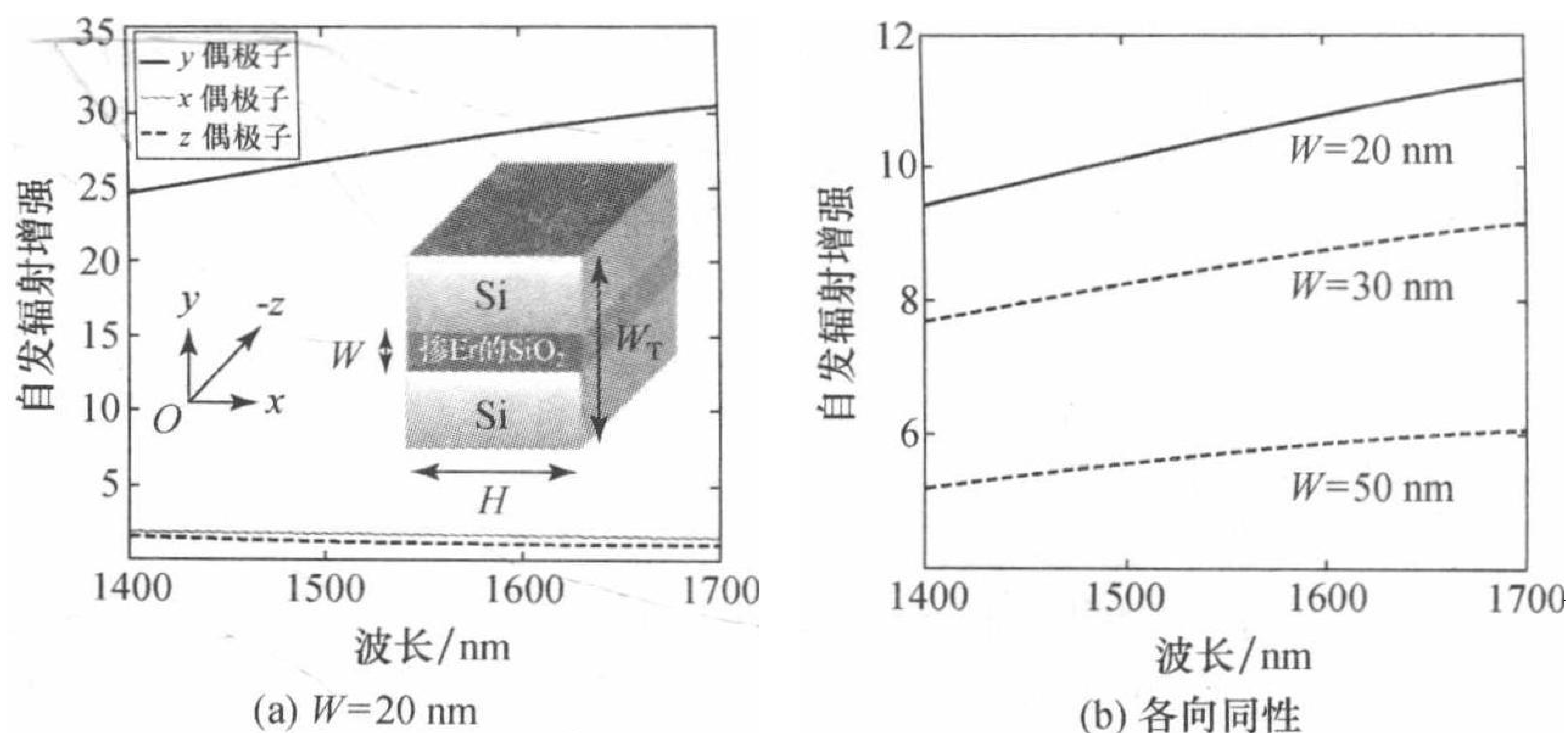


图 5.4 槽形波导结构增强光发射

5.3.3 锗硅材料

锗的直接带隙只比间接带隙高 136 meV，早在 1978 年人们就观察到了低温下高功率密度激光激发产生的锗直接带隙发光. 近年来，美国麻省理工学院 Kimerling 研究组[32]提出了通过能带工程获得外延锗光源的理论，并获得了令人瞩目的实验结果. 这一原理为，在硅基底上外延生长的锗薄膜由于热膨胀系数

失配，退火后锗薄膜内产生了张应变，使能带结构发生变化. 0.25%的张应变会使 Γ 与 L 谷的带隙差减小到 115 meV，这一差别可以通过 n 型掺杂来进一步弥补. 10^{19} 的 n 掺杂可以填满 L 谷并使电子开始填充 Γ 谷，从而获得直接带隙跃迁发光和显著的增益，如图 5.5 所示.

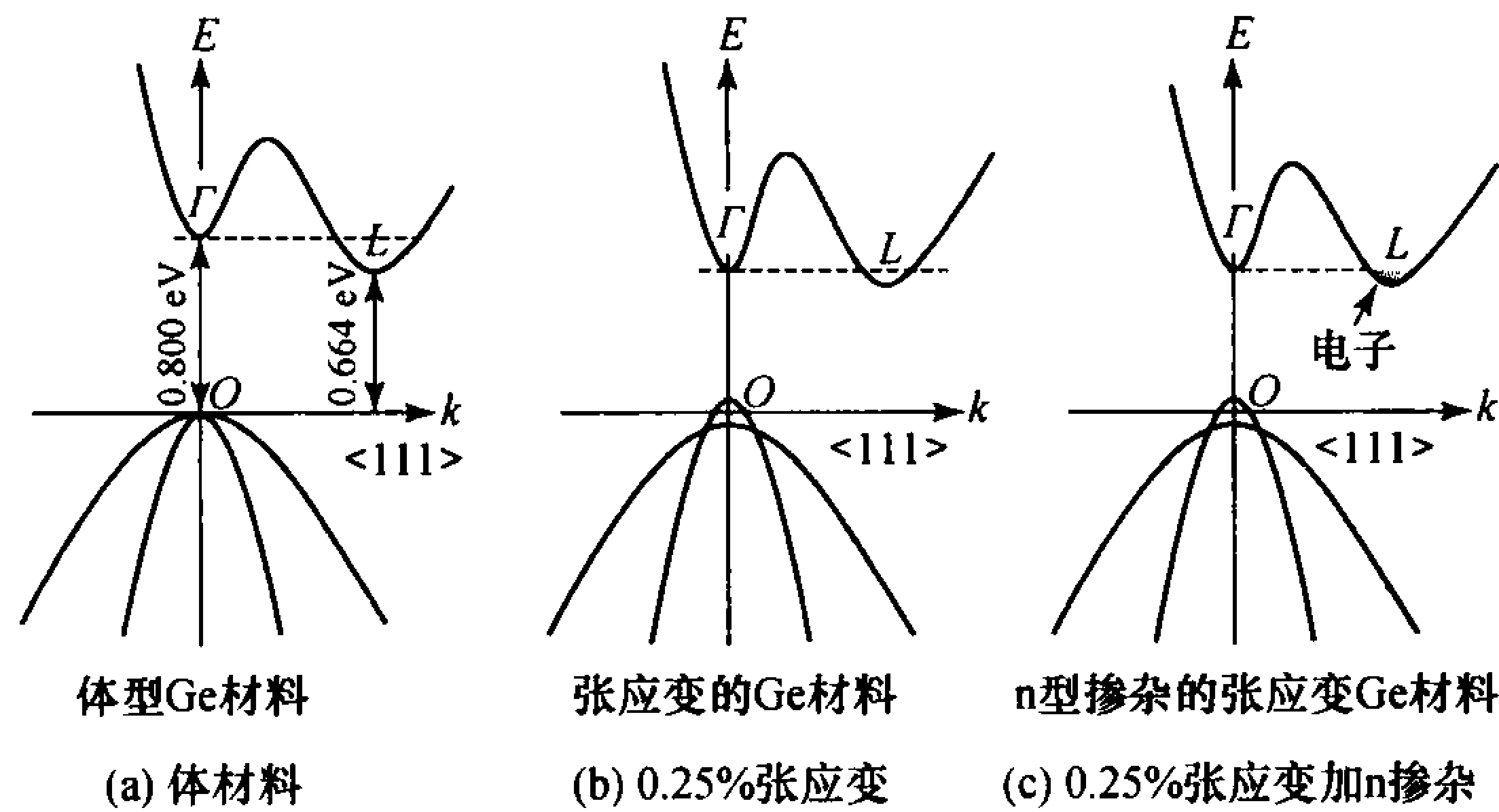


图 5.5　锗的带隙结构

§5.4　硅基发光二极管

5.4.1　发光二极管的工作特性[1~3]

发光二极管(LED)为半导体光电子器件，它的基本特性包括电学特性、光学特性和光电转换特性. 电学特性主要体现在二极管上电流随电压的变化关系；光学特性主要体现其发光强度、光谱的峰值波长、光谱半高宽；光电转换效率反映器件的量子效率，能量转换效率；等等.

LED 的主要光学参数包括光通量、发光效率、发光强度、光强分布、发光波长、光谱特性和温度特性等等.

(1) 发光效率和光通量. 发光效率微观上是指半导体的体内复合产生的光子数同注入的电子-空穴对数之比；宏观上就是光通量与电功率之比. 发光效率表征了光源的节能特性，这是衡量现代光源性能的一个重要指标.

(2) 发光强度和光强分布. LED 发光强度是表征它在某个方向上的发光强弱，定义为在正向工作电流驱动下发出的光强(单位为 cd). 由于 LED 在不同的空间角度光强相差很多，因此 LED 的光强分布特性这个参数实际意义很大，直接影响到 LED 显示装置的最小观察角度. 比如体育场馆的 LED 大型彩色显示屏，如果选用的 LED 单管分布范围很窄，那么面对显示屏处于较大角度的观众将看到失真的图像.

(3) 发光波长. 对于 LED 的光谱特性我们主要看它的单色性是否优良，而且要注意到红、黄、蓝、绿、白色 LED 等主要的颜色是否纯正. 因为在许多场合下，比如交通信号灯对颜色就要求比较严格，不过据观察现在我国的一些 LED 信号灯中绿色发蓝，红色的为深红，从这个现象来看我们对 LED 的光谱特性进行专门研究是非常必要而且很有意义的.

(4) 光谱特性. 对于直接带隙半导体，峰值波长与禁带宽度相对应，也就是只能对大于禁带宽度的光子能量做出响应. 由于不同材料有不同的禁带宽度，因此光谱响应范围也不相同. 硅发光二极管主要工作范围在 1.1 μm 以下. 锗发光二极管主要在 1.6 μm 以下. 对于间接带隙半导体，峰值波长由等电子陷阱发光中心的位置决定. 在半导体中，参与电子-空穴复合的能带有一定宽度，而不是能级之间的载流子复合发光，因此导带底附近和价带顶附近的能态都会对发光有贡献，造成了发光二极管的发射光谱较宽. 通常光谱半高宽为 50～100 nm.

(5) 温度特性. 半导体材料对温度非常敏感，发光波长和发光强度都随着温度的变化而变化，这是因为半导体的禁带宽度随温度的变化而变化，如磷砷化镓发光二极管是在半导体中掺入特定的杂质形成 pn 结，当这种二极管加正向电压并导通电流足够大时，n 区和 p 区的多数载流子向对方扩散，并与对方的多数载流子复合释放能量发光，它导电是和内部 pn 结的自由电子和空穴有关的，而这些又与温度有关，所以在潮湿的环境中，会影响 pn 结多数载流子的分布，从而影响加电压后多数载流子的移动的复合，进而影响发光波长和强度.

5.4.2 体硅发光二极管

作为硅基光电子学中的光源，基于带间光学跃迁自发辐射的硅基发光二极管的实现是硅基光子学研究中的一个主攻方向. 目前的研究重点有两个，一是如何采用适宜的有源区材料，实现其高效率和高稳定度的发光. 二是从器件实用化角度考虑，如何实现硅 LED 在室温下的电致发光. 迄今，人们已尝试了三种硅基纳米材料用于高效率硅 LED 的制作，即高纯单晶体硅，硅纳米量子点和掺 Er 硅.

单晶体硅是一种典型的间接带隙半导体材料，它一直被人们认为是不发光的. 但是近年的研究指出，对体单晶硅进行有效的材料改性，也可以获得高效率的发光. 澳大利亚的 Green 研究小组[35]在《自然》上报道了他们在块体硅发光方面的研究结果，以区熔法生长的具有晶格完整性好的硅单晶作为基底，利用适宜的蚀刻技术使表面加以构型，把硅表面设计成锯齿状光学图形，使入射角小于全反射角，光的输出效率可以达到 99%以上，如图 5.6 所示. 并对其进行高质量的热氧化和表面钝化，以有效地减少载流子的非辐射复合速率，可使量子效率得以明显提高，其室温下的电注入有效量子效率可达 1%以上，开启电压小于

1V. 利用将高剂量的硼离子注入 n 型硅单晶中所形成的位错环势垒所具有的二维量子限制作用，也已在室温下观测到了强电致发光.

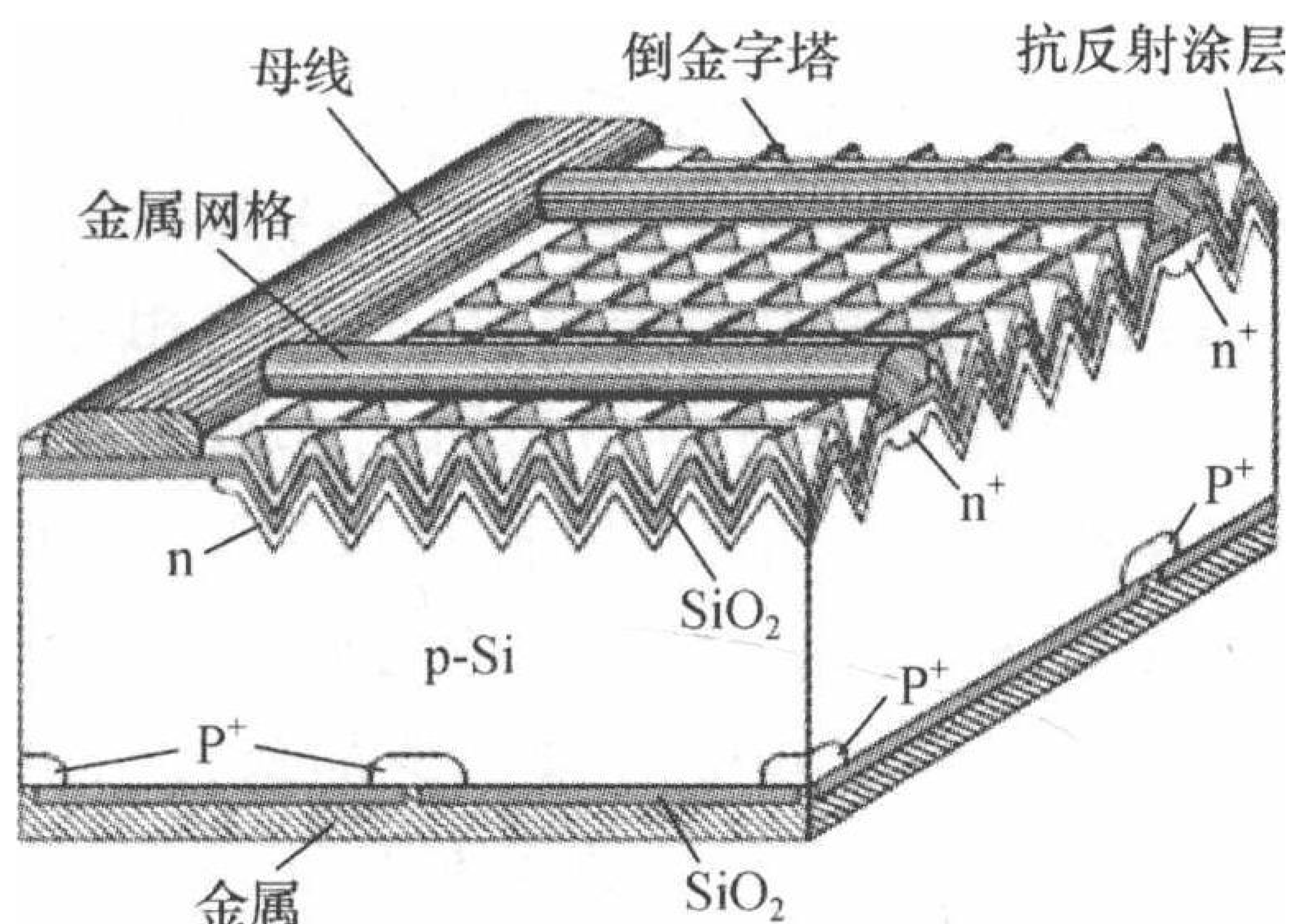


图 5.6 金字塔形体硅发光二极管的示意图

5.4.3 纳米硅发光二极管

硅纳米量子点或纳米晶粒是一种具有强三维量子限制效应的典型硅基纳米结构，在硅 LED 的制作中具有潜在优势. 富硅氧化硅中的纳米硅因为量子限制效应可以增加载流子的辐射跃迁几率，提高电子-空穴的注入效率，另外还可以通过控制纳米晶的尺寸控制受激发射的波长. 采用这种硅基低维材料制作的硅-LED，已获得从红外到紫外光波长范围的强光发射. 例如，由镶嵌在 SiN_x 膜层中的硅纳米晶粒所制成的电致发光 LED，室温下的外量子效率可高达 1.6%[36]. 如果将硅纳米晶粒置入一个具有强光限制效应的微腔结构中，还可以实现某一波长范围的窄谱线强光发射[37]. 尤其是未来的硅基蓝光发射器件，很有可能在仅具有 1～2 nm 的超小硅纳米团簇中实现. Priolo 研究小组[38]采用等离子体化学气相沉积的方法制备不同纳米硅晶浓度的富硅氧化硅材料，并制备出 LED，在 850 nm 观察到了很强的电致发光. 富硅氧化硅材料发光特性和纳米硅晶的浓度、尺寸以及退火温度有很强的关系，而控制纳米硅晶的浓度和尺寸的工艺比较复杂，目前还没有实用化.

5.4.4 铒掺杂硅和铒硅酸盐化合物发光二极管

铒(Er)是一种有效的发光中心，采用镶嵌于 SiO_2 层中掺 Er 的硅纳米晶粒(掺 Er 富硅氧化硅)可以获得 1.54 μm 波长的强光发射. 利用等离子体化学气相沉积形成的掺 Er 富硅氧化硅薄膜所试制的 MOS 型硅 LED 结构，已经实现了量子效率高达 1%的 1.54 μm 强光发射[39]. 最近，北京大学秦国刚院士研究组将 Er 的硅酸盐化合物材料制成 LED，观察到了光通信波段 1.53 μm 电致发光[20].

§5.5 硅基激光器

硅基激光器的研制是硅基光电子学领域中的一个最具有魅力、最富挑战性的前沿课题. 说最具有魅力，是因为它有着丰富而深刻的物理内涵；说最富有挑战性，因为它作为硅基光电子集成中的光源，研究具有相当大的难度. 按照光的受激辐射原理，要设计并制作出具有器件实用化水平的硅基激光器，需要具备下述几个条件：① 制备出具有光增益、光放大和受激辐射的有源区材料或结构；② 能够实现粒子数的反转；③ 具有适宜结构形式的光学谐振腔；④ 能够实现电注入条件下的受激辐射，而这其中的每一步都要付出艰苦的探索. 目前，人们已初步提出了几种能产生光增益或受激辐射的方案，即具有高密度和小尺寸的有序硅纳米晶粒、受激拉曼散射、GeSi 激光器、Ⅲ-Ⅴ族-硅混合激光器.

5.5.1 纳米结构硅激光器

2005 年，Cloutier 等[40]在 SOI 的晶体硅上刻蚀了一种周期性的纳米孔阵列结构，孔直径为 110 nm. 在连续 1.5 W 的 514.5 nm Ar 离子激光泵浦下获得了 1.278 μm 的连续光. 这是由于周期性的纳米孔阵列产生了高密度的 A 型陷阱中心，这些缺陷中心作为光激活中心发生反转产生受激发射和光增益. 但主要缺点是只有在低于 80K 的温度才能产生受激发射，如图 5.7 所示.

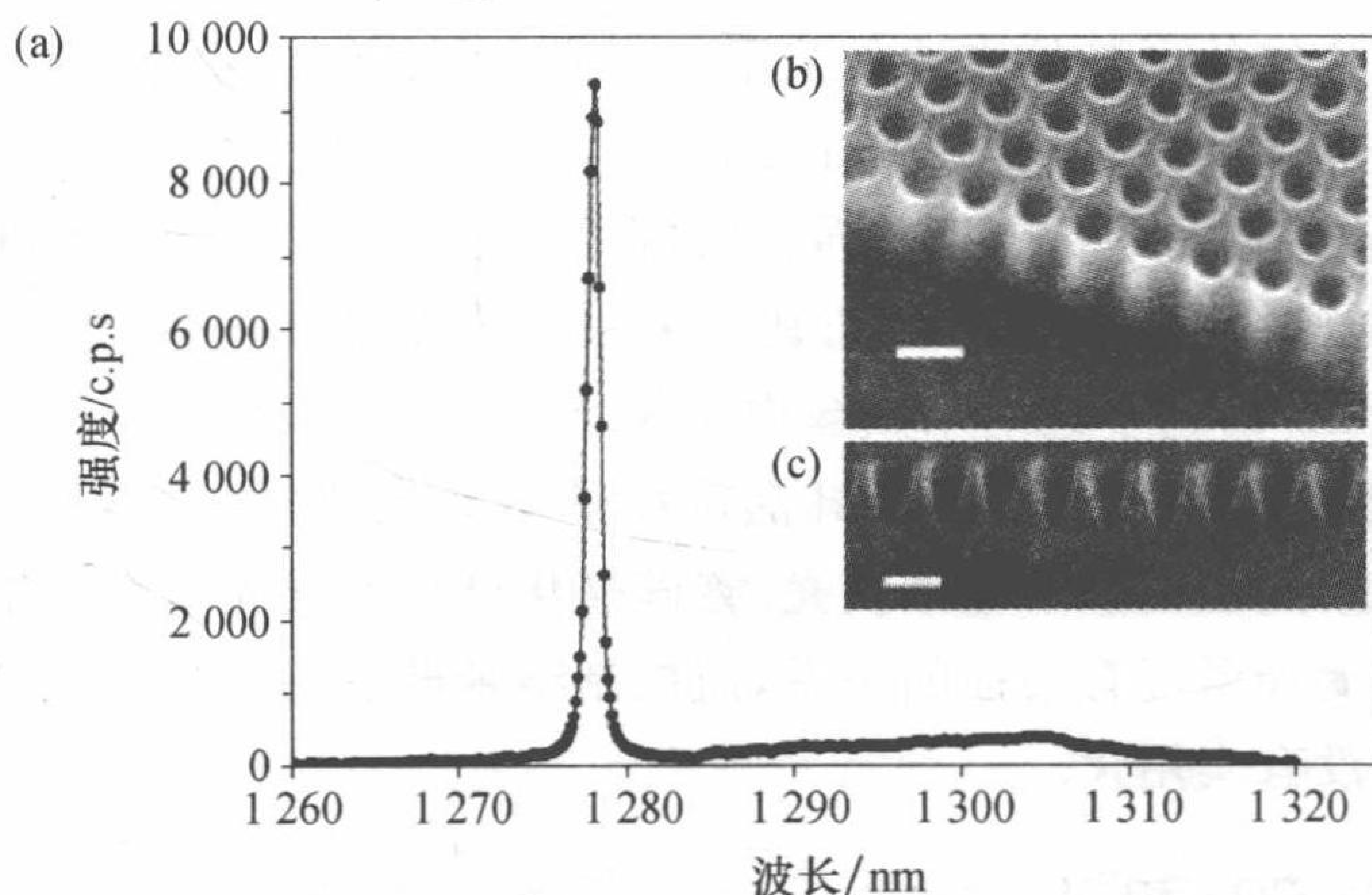


图 5.7 光致激光谱(a)以及纳米结构硅激光器的表面(b)、端面(c)SEM 照片

5.5.2 硅拉曼激光器

拉曼散射是光通过介质时入射光与分子运动相互作用而引起散射光频率发生变化(位移也相应发生变化)的散射. 拉曼散射中，不仅光子的动量要发生改

变，其能量也要发生改变，其原因是入射光子与分子碰撞时，入射光子与分子交换能量，使分子的振转能级发生改变，从而散射出不同能量的光子，散射光子能量低于入射光子能量的散射叫做斯托克斯散射，散射光子能量高于入射光子能量的散射叫做反斯托克斯散射. 硅材料中拉曼散射是一种非线性光学现象，只有高功率密度的光入射硅材料才会出现较强的拉曼散射光，在此过程中高能量密度的泵浦光被硅吸收并转换为激光能量，同时硅材料完成振转能级间的跃迁，这样就克服了硅材料发光效率低的缺点.

2005 年，Intel 公司的研究人员 Rong 等[41]研制成功了第一台全硅拉曼激光器，其硅波导为 S 形结构，总长度 $D=4.8$ cm，波导宽度 $W=1.5$ μm、高 $H=1.55$ μm、刻蚀深度 $h=0.7$ μm，如图 5.8 所示. 将硅波导制成 S 型可以减少硅片尺寸，便于硅器件的集成. 采用 p-i-n 二极管结构，通过在硅波导两侧加反向偏置电压，来减少波导内的载流子浓度，从而抑制自由载流子对激光的吸收. 当 p-i-n 结两侧反向偏置电压为 25V 时，激光阈值为 0.4 m，斜率效率为 9.4%. 紧接着，该小组又通过对器件结构进行合理改善，利用由多层介电薄膜覆盖的 SOI 光波导形成的光学微腔，实现了稳定的单模激射输出，并研制成了连续拉曼激光器[42]，2008 年又实现了级联拉曼激光器. 但该拉曼激光器需要很高功率的光泵浦，从原理上无法实现电泵浦的拉曼激光器，限制了其进一步的发展.

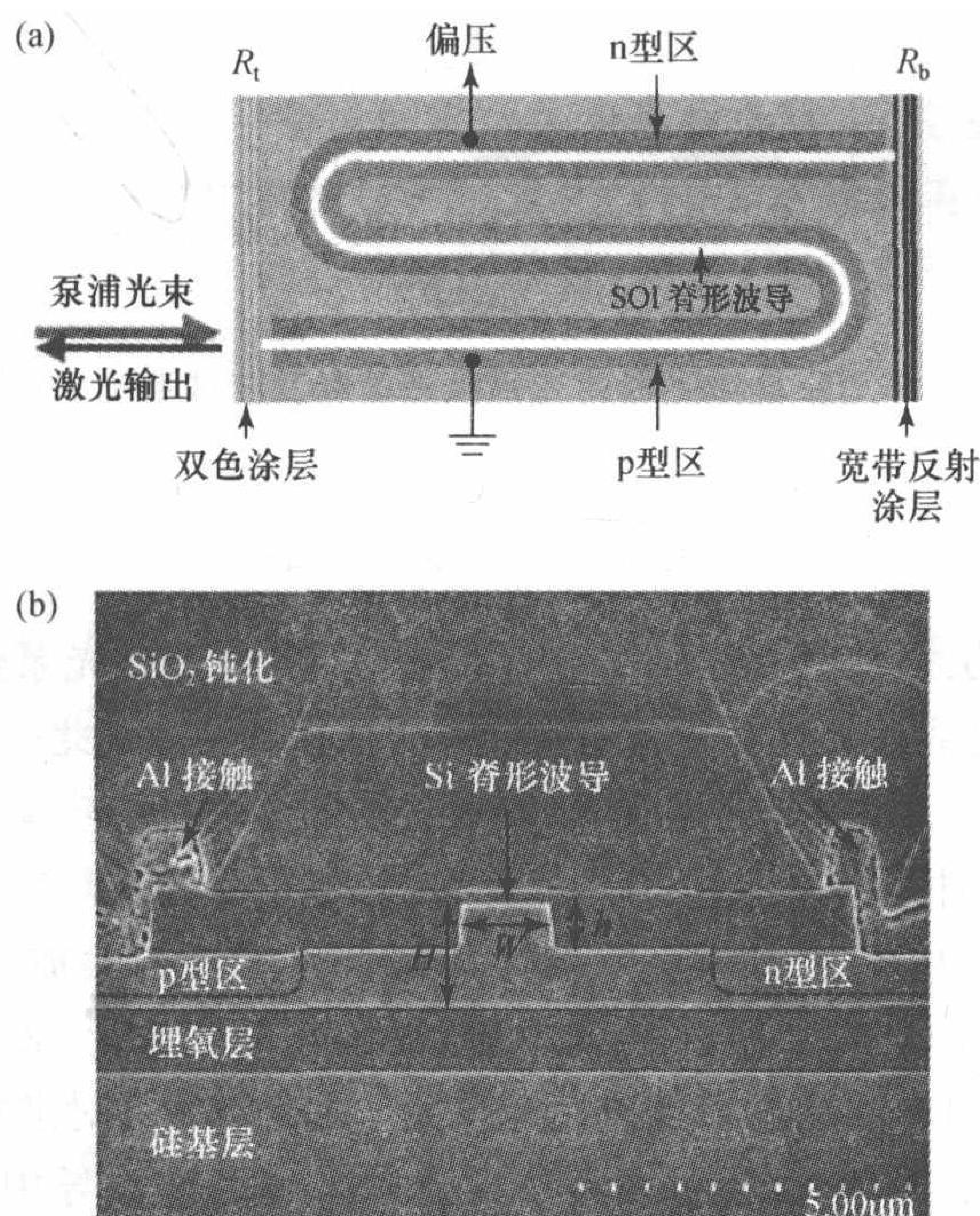


图 5.8 硅拉曼激光器的结构示意图(a)和端面的 SEM 照片(b)

5.5.3 锗硅激光器

2010 年,麻省理工学院的 Kimeriling 研究组[43]在前期硅基高响应率高速的外延锗探测器和低功耗外延 Ge 光电调制器的基础上[44],研制了世界上首个光通信波段室温工作的 Ge 激光器,如图 5.9 所示.其基本原理是基于拉应力和 n 型重掺杂拉平 Ge 的直接带隙和间接带隙方法,尽管锗属于间接带隙的半导体,复合发光需要声子参与导致效率低,但通过能带工程把间接带隙变成直接带隙实现激光,制备出光泵浦的 Ge 激光器.正如Ⅲ-Ⅴ族激光器研究发展的历程一样,电注入激射将成为下一个 Ge 硅激光器攻克的目标.令人惊喜的是,2012 年 Kimeriling 研究小组又成功实现了室温电致 Ge 激光器[45].

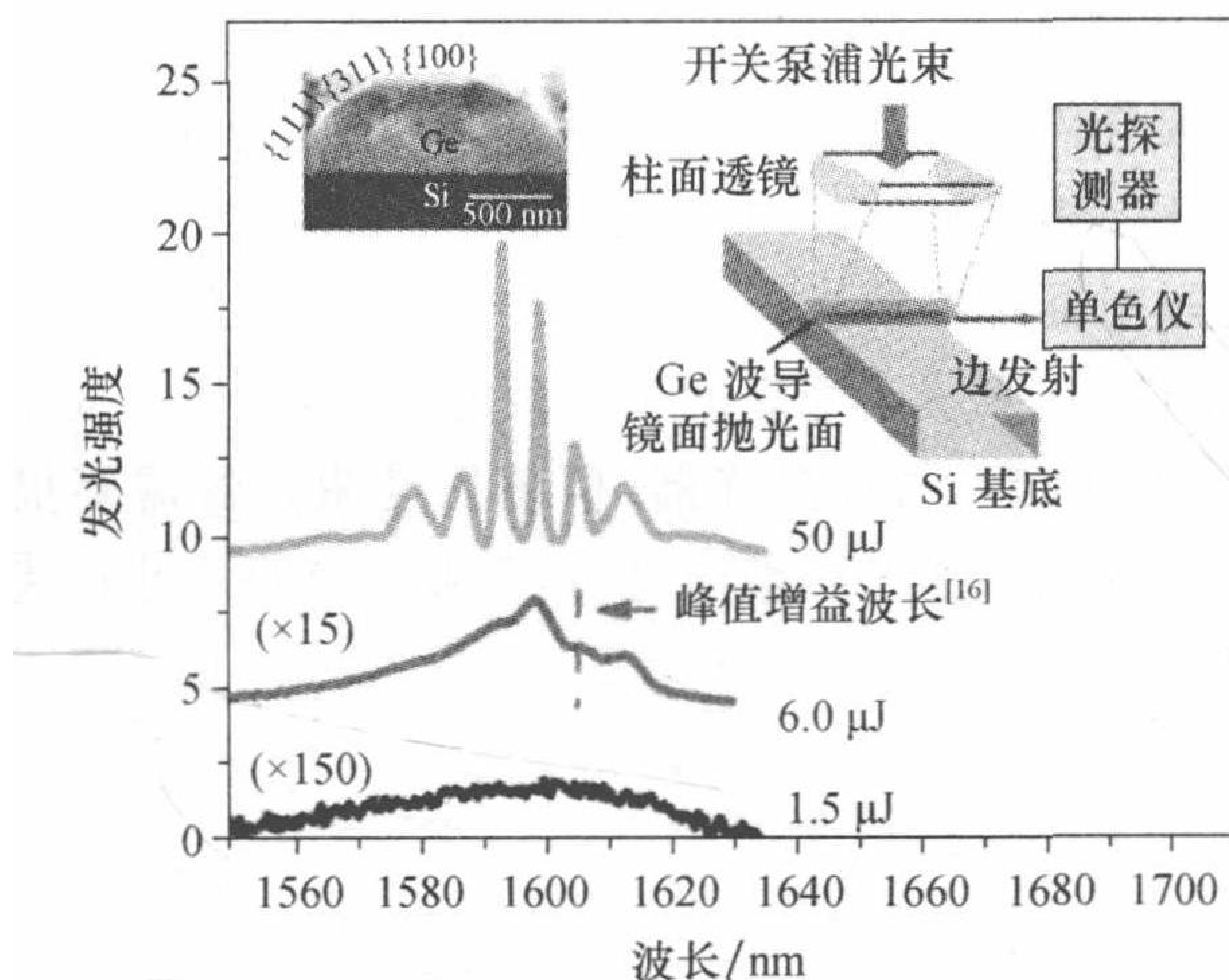


图 5.9 Ge 硅激光器的光致激光谱以及测试系统示意图

5.5.4 Ⅲ-Ⅴ族化合物和硅混合集成激光器

上面几种方法虽然都观察到了一些激射现象,但都是光泵浦的,并且泵浦功率都很大,这和硅基片上光源电致、低功率的要求不符.因此一种变通的方法是采用混合集成的片上光源,利用Ⅲ-Ⅴ族发光性能好的优点,采用键合等工艺将Ⅲ-Ⅴ族材料的器件与硅基光电子器件混合集成在同一个基底上.

图 5.10 所示为 Bower 小组[46]设计并制备出的混合集成的激光器.利用键合的方法将增益介质Ⅲ-Ⅴ族材料(InP)制备在硅基基底上,器件结构的制备以及与下层硅波导的对准依靠普通光刻来实现,具有较高的对准精度.利用电注入在 InP 材料产生激光,然后利用倏逝波耦合进下层的硅波导中.图 5.10 所示为器件横截面结构示意图.该激光器还可以同时在一个硅基底上集成多个激光器.

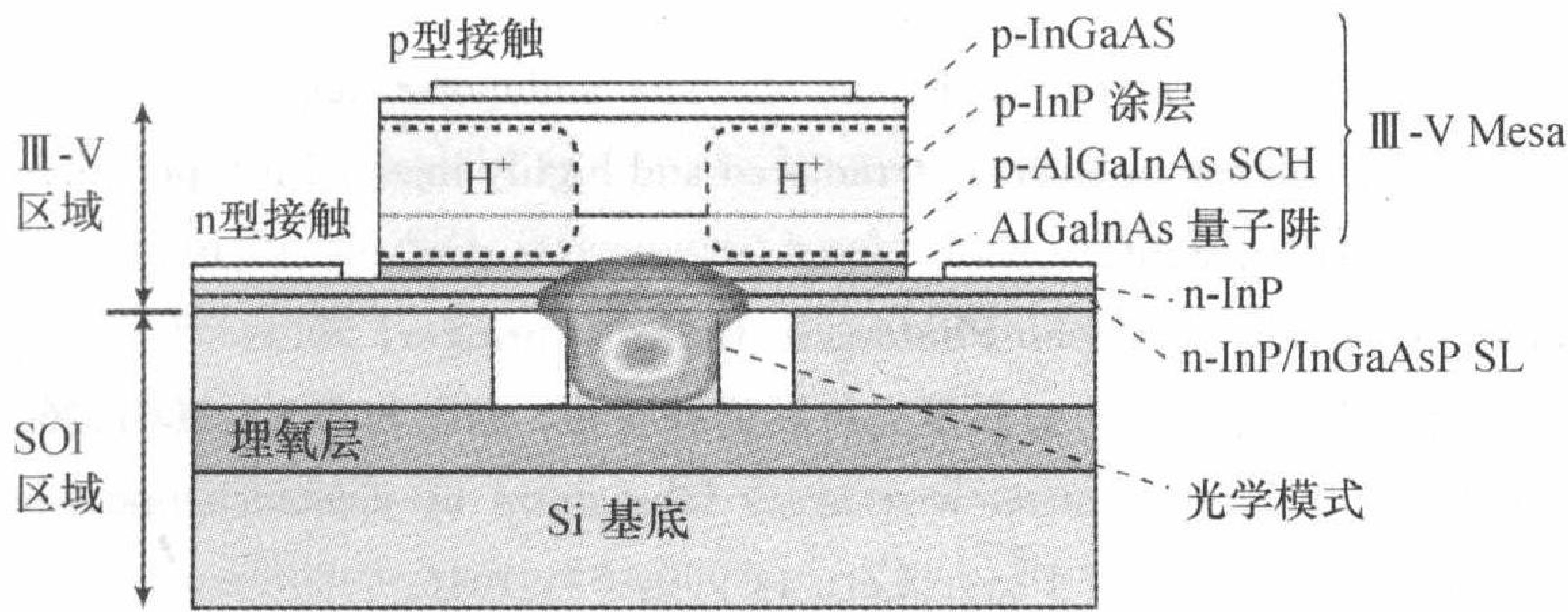


图 5.10 Ⅲ-V 族-Si 混合激光器的端面示意图

§5.6 总 结

虽然硅基光源近年来取得了很大进展，但所面临的挑战仍是多方面的. 首先，外部光源和混合集成的片上光源的优点是具有高的发光效率，容易实现连续或者可调的激光输出. 但其缺点是集成度不高，制备工艺复杂，与 CMOS 工艺不兼容，系统成本昂贵，还必须考虑到输出激光到波导的耦合对准问题；第二，单片集成硅基光源的优点是集成度高，成本低，但目前最大的困难是发光效率低，电致发光困难. 其中，在掺铒硅光源所面临的困难是：如何克服自由载流子吸收实现净增益，这即需要深入的了解电激发的能量传输机制，又需要仔细的设计器件结构；锗作为间接带隙材料，其直接带隙发光与Ⅲ-V 族材料有很大不同，这些特点对制造激光器所可能产生的影响尚不清楚，再者，锗的高折射率(4.3)对输出光的耦合是一个很大的挑战.

综合分析各种硅基光源的优缺点和近年来国际上研究动向，我们可以看到，硅基材料掺铒仍然是重要的研究方向，槽型波导是电注入激光器的一种有前景的结构，而基于槽型波导的宽带发光增强效应则是另一个较新的物理现象，其潜在的影响尚不清晰. 另外，锗电致光源也是能够尽快取得突破的途径之一，一旦锗电致激光器得以实现，硅基光电子学必定会进入一个高速发展期. 当然Ⅲ-V 族化合物和硅混合集成激光器是最接近实用的一个方向，各国也都在加大力度攻克它的一些实用前的技术问题，期望很快接近实用.

参考文献

[1] 余金中. 半导体光电子技术. 北京：化学工业出版社，2003.
[2] 张季熊. 光电子学教程. 广州：华南理工大学出版社，2009.
[3] 彭英才，赵新为，傅广生. 硅基纳米光电子技术. 石家庄：河北大学出版社，2009.
[4] Pavesi L. Silicon-based light sources for silicon integrated circuits. Adv. Opt. Tech.,

2008, 2008: 416926.

[5] Jonsson P, Bleichner H, Isberg M, et al. The ambipolar auger coefficient: measured temperature dependence in electron irradiated and highly injected n-type silicon. J. Appl. Phys., 1997, 81(5): 2256.

[6] Pavesi L, Lockwood D. Silicon Photonics. Berlin: Springer, 2004.

[7] 曾刚，杨宏春，阮成礼，等. 硅基发光材料研究进展. 压电与声光，2004,26: 296.

[8] Chanham L T. Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers. Appl. Phys. Lett., 1990, 57: 1046.

[9] Lehman V, Gosele U. Porous silicon formation: a quantum wire effect. Appl. Phys. Lett., 1991, 58: 856.

[10] Ssnders G D, Chang Y C. Theory of optical properties of quantum wires in porous silicon. Phys. Rev. B., 1992, 45: 209.

[11] Kanemitsu Y. Quantum-size effects of interacting electrons and holes in semiconductor microcrystals with spherical shape. Phys. Rev., 1988, B38: 9797.

[12] 朱美芳，陈国，许怀哲，等. 镶嵌在氧化硅薄膜中纳米晶硅的可见光荧光谱与发光机制的研究. 物理学报，1997, 46: 1645.

[13] Qin G G. Mechanism of the visible luminescence in porous silicon. Solid State Communications, 1993,86(9): 559.

[14] Kik P G, Polman A. Gain limiting processes in Er-doped Si nanocrystal waveguides in SiO_2. J. Appl. Phys., 2002, 91(1): 534.

[15] Shin J L J, Park N. Optical gain in Si-nanocrystal sensitized, Er-doped silica waveguide using top-pumping 470 nm LED. Proc. Opt. Fiber Communication Conference, 2004.

[16] Toccafondo V, Pasquale F D, Faralli S, et al. Study of an efficient longitudinal multimode pumping scheme for Si-nc sensitized EDWAs. Opt. Exp., 2007, 15(22): 14907.

[17] Isshiki H, de Dood M J A, Kimura T. Self-assembled infrared-luminescent Er-Si-O crystallites on silicon. Appl. Phys. Lett., 2004, 85: 4343.

[18] Savio R L, Miritello M, Piro A M, et al. The influence of stoichiometry on the structural stability and on the optical emission of erbium silicate thin films. Appl. Phys. Lett., 2008, 93(2): 1919.

[19] Wang X J, Kimura T, Zhou Z. Enhanced Er^{3+} luminescence of Er Silicate by Y and Yb co-doping. Proc. 6th IEEE International Conference on Group IV Photonics, 2009.

[20] Yin Y, Sun K, Xu W J, et al. 1.53 μm photo- and electroluminescence from Er^{3+} in erbium silicate. J. Phys.: Condens. Matter, 2009, 21(1): 12204.

[21] Wang X J, Nakajima T, Kimura T. Fabrication and characterization of Er silicates on SiO_2/Si substrates. Appl. Phys. Lett., 2009, 95(4): 1906.

[22] Wang X J, Wang B, Wang L, et al. Extraordinary infrared photoluminescence efficiency of $Er_{0.1}Yb_{1.9}SiO_5$ films on SiO_2/Si substrates. Appl. Phys. Lett., 2011,98(7): 1903.

[23] Han H S, Seo S Y, Shin J H, et al. Coefficient determination related to optical gain in

erbium-doped silicon-rich silicon oxide waveguide amplifier. Appl. Phys. Lett., 2002, 81: 3720.

[24] Wang X J, Yuan G, Isshiki H, et al. Photoluminescence enhancement and high gain amplification of $Er_xY_{2-x}SiO_5$ waveguide. Appl. Phys., 2010, 108(1): 3506.

[25] Guo R M, Wang X J, Zhou Z. Optical amplification in Er/Yb silicate strip loaded waveguide. Appl. Phys. lett., 2011, 99(16): 1115.

[26] Kalkman J, Strohhfer C, Gralak B. Surface plasmon polariton modified emission of erbium in a metallodielectric grating. Appl. Phys. Lett., 2003, 83: 30.

[27] Wang Y, Zhou Z. Strong enhancement of erbium ion emission by a metallic double-grating. Appl. Phys. Lett., 2006, 89: 18.

[28] Polman A, Min B, Kalkman J, et al. Ultralow-threshold erbium-implanted toroidal microlaser on silicon. Appl. Phys. Lett., 2004, 84(7): 1037.

[29] Lu T, Yang L, Loon R V, et al. On-chip green silica upconversion microlaser. Opt. Lett., 2009, 34: 482.

[30] Makarova M, Sih V, Warga J, et al. Enhanced light emission in photonic crystal nanocavities with Erbium-doped silicon nanocrystals. Appl. Phys. Lett., 2008, 92 (16): 1107.

[31] Jun Y C, Briggs R M, Atwater H A, et al. Broadband enhancement of light emission in silicon slot waveguides. Opt. Exp., 2009, 17: 7479.

[32] Liu J F, Sun X C, Pan D, et al. Tensile-strained, n-type Ge as a gain medium for monolithic laser integration on Si. Opt. Exp., 2007, 15(18): 11272.

[33] Sun X C, Liu J F, Kimerling L C, et al. Direct gap photoluminescence of n-type tensile-strained Ge-on-Si. Appl. Phys. Lett., 2009, 95(1): 1911.

[34] Sun X C, Liu J F, Kimerling L C, et al. Room-temperature direct bandgap electroluminescence from Ge-on-Si light-emitting diodes. Opt. Lett., 2009, 34(8): 1198.

[35] Green M A, Zhao J, Wang A, et al. Efficient silicon light-emitting diodes. Nature, 2001, 412(6849): 805.

[36] Chok S, Park N M, Kim T Y. High efficiency visible elecroluminescence from silicon nanocyrstals embedded in silicon nitride using a transparent doping layer. Appl. Phys. Lett., 2005, 76(7): 1909.

[37] Toshikiyo K, Fujii M, Mayashi S. Enhanced optical properties of Si nanocyrstal in planar microcavity. Physica E., 2003, 17: 451.

[38] Pavesi L, Negro L D, Mazzoleni C, et al. Optical gain in silicon nanocrytals. Nature, 2000, 408: 440.

[39] Franz'o G, Coffa S, Priolo F, et al. Mechanism and performance of forward and reverse bias electroluminescence at 1.54 μm from Er-doped Si diodes. J. Appl. Phys., 1997, 81 (6): 2784.

[40] Cloutier S G, Kossyrev P A, and Xu J. Optical gain and stimulated emission in periodic nanopatterned crystalline silicon. Nature Materials, 2005,4(12): 887.

[41] Rong H, Liu A, Jones R, et al. An all-silicon Raman laser. Nature, 2005,433: 292.

[42] Rong H, Jones R, Liu A, et al. A continuous-wave Raman silicon laser. Nature, 2005, 433: 725.

[43] Liu J F, Sun X C, Aguilera R C, et al. Ge-on-Si laser operating at room temperature. Opt. Lett., 2010,35(5): 679.

[44] Liu J F, Sun X C, Kimerling L C, et al. Direct gap optical gain of Ge-on-Si at room temperature. Opt. Lett.,2009,34: 1738.

[45] Camacho-Aguilera R E, Cai Y, Patel N, et al. An electrically pumped germanium laser. Opt. Exp., 2012,20(11): 3117.

[46] Koch B R, Fang A W, Chang H H, et al. A 40 GHz Mode Locked Silicon Evanescent Laser. Proc. 4th International Conference on Group IV Photonics, 2007.

第六章　硅基光学调制

光学调制，利用其他能量形式改变光信号的特征，是集成光路中的一个重要环节. 传统的光学调制通常采用铌酸锂、Ⅲ-Ⅴ族半导体、或其他材料作为载体，其主要原因是因为这些材料的线性电光效应(Pockel 效应)非常显著，易于实现高速调制. 然而用这些材料制作的器件体积大，成本高，而且与硅基微电子工艺不兼容，无法进行大规模生产. 近年来，硅基电光调制在理论和技术获得巨大突破，其调制速率已经从 MHz 进入到 GHz 范畴. 硅基高速调制不仅是未来光交叉连接(optical cross-connect, OXC)和光分插复用器(optical add-drop multiplexer, OADM)系统中的核心技术，而且在芯片光互连和光计算技术中也具有很大的应用前景. 因此，对硅基光学调制进行系统的研究意义重大.

本章从一般性的光学调制入手，系统阐述了几种硅基光学调制的原理和机制，给出了调制器件设计的重要考查指标，并结合几个具体调制器的设计对物理原理如何在器件中实现给予了阐释，并由此对进行新型调制器设计给出了一些提示.

§6.1　光学调制原理

光学信号可以由强度、振幅、频率、相位、偏振及传播方向等特征参数来表征. 光学调制则是改变光学信号的一个或多个特征参数的过程. 由于人的眼睛以及大部分的光学探测器对光的强度(振幅)非常敏感，而且其他光学参数(如频率、相位、偏振等)的变化也都可以通过强度(振幅)的变化来表达，因此，光学调制的最终效果是通过光的强度(振幅)的变化来检测的.

正如光在材料中的行为可以通过材料的折射率变化来预测一样，光学调制的过程实际上也是一个材料折射率变化的过程. 光学调制的原理也就是利用各种能量形式使材料折射率发生变化的理论. 下面简单介绍一下电光、热光、声光调制.

6.1.1　电光调制

电光效应指的是外加电场引起的材料折射率的变化，也就是电场对材料光学性质的调制—电光调制. 外加电场通过改变原子或分子中电子的运动规律，或者材料的晶体结构，使得原本各向同性的材料变成各向异性，并由此产生材料的

双折射效应.这种材料折射率对外加电场的依赖关系可以表示为

$$n' = n'(E), \tag{6.1.1}$$

其泰勒展开式可以表述为

$$n' = n + a_1 E + a_2 E^2 + a_3 E^3 + \cdots, \tag{6.1.2}$$

其中

$$\Delta n = a_1 E_1 \tag{6.1.3}$$

和

$$\Delta n = a_2 E^2 = (\lambda K) E^2 \tag{6.1.4}$$

被分别称为泡克耳斯效应和克尔效应.在泡克耳斯效应中,外加电场引起的晶体折射率变化正比于电场强度,因此又称做线性电光效应.而克尔效应也被称为二阶非线性电光效应,因为外加电场引起的晶体折射率变化正比于电场强度的平方.一般来说,材料的克尔效应强度远低于泡克耳斯效应强度.

从式(6.1.3)和(6.1.4)可以得到:① 对于非晶材料和具有对称性的晶体,其介电特性是各向同性的,因此,由正电场和负电场引起的折射率变化相等,这意味着:$\Delta n = a_1 E = a_1(-E)$,从而,$a_1 = 0$;对这些材料,不存在泡克耳斯效应;② 对于那些具有非对称中心的晶体,泡克耳斯效应是存在的;③ 无论外加电场是正还是负,a_2 总不等于 0;这表明无论对什么材料,只要有外加电场,就会有克尔效应产生.

6.1.2 热光调制

对应于材料所处的温度不同,其分子或晶体结构发生变化,从而造成材料的光学性质不同.这种物理现象称为热光效应(thermo-optic effect),可以用下式表达:

$$n(T) = n_0 + \alpha(T), \tag{6.1.5}$$

其中 n_0 是该材料在某一特定温度下的折射率,α 是该材料的热光系数,T 表示温度.显然,通过人为外加的能量形式可以改变材料的温度,从而改变材料的光学性质.这个过程就叫做热光调制.

6.1.3 声光调制

声光效应是弹性波与光波的相互作用所表现出来的一种物理现象.当超声波通过介质时会造成介质的局部压缩和伸长而产生弹性应变,而且该应变随时间和空间作周期性变化,从而使介质的折射率发生改变.当光通过这一受到超声波扰动的介质时就会发生衍射现象,这种现象称之为声光效应(acousto-optic effect)[1].

存在于超声波中的此类介质可视为一种由声波形成的位相光栅(称为声光

栅)，其光栅的栅距(光栅常数)即为声波波长. 当一束平行光束通过声光介质时，光波就会被该声光栅所衍射而改变光的传播方向，并使光强在空间作重新分布. 在声光效应中，弹性应变引起介质折射率的变化与声功率的关系为[2]

$$\Delta n = \sqrt{n^6 p^2 10^7 P_a/(2Qv_a^3 A)}, \tag{6.1.6}$$

其中 n 是无应力时介质的折射率，p 是介质的弹光张量，P_a 是超声驱动总功率，Q 是质量密度，v_a 是声速，A 是声波场与介质的交叠区域面积，式中全部采用标准单位. 定义一个声光系数

$$M_2 = n^6 p^2/(Qv_a^3), \tag{6.1.7}$$

则表达式简化为如下形式：

$$\Delta n = \sqrt{M_2 10^7 P_a/(2A)}. \tag{6.1.8}$$

在晶体中，声光效应主要是由晶体取向决定，也就是与 p 相关. 但是即使选择最优的材料和取向，这种效应相对来说也较小. 例如：光波长为 6328Å，石英的 M_2 为 1.51×10^{-18} s³/(g·m)，而 $LiNbO_3$ 的 M_2 为 6.9×10^{-18} s³/(g·m)，即使声功率密度为 100 W/cm²，折射率的改变也只在 10^{-4} 数量级. 虽然折射率改变较小，但每一个声波峰值对光都会起作用，如果相位条件满足，作用的结果积累起来就可形成衍射效应.

§6.2　光学调制评价

在光学调制的实际作用中，需要对调制方法以及执行调制的器件进行评价. 以下是利用各种效应进行调制器设计时都必须尽量优化的性能指标.

1. 调制带宽

这一指标衡量调制器往光波上加载数据能力的大小. 是调制器最重要的性能指标之一. 通常可以定义为透射率减小到最大值的 50% 的变化频率. 目前，硅基光学调制器带宽约为几十 GHz 数量级，已经达到商用非集成光学调制器的水平.

2. 调制深度

这一指标为调制器工作时透射率的最大值 I_{max} 与最小值 I_{min} 的比率. 通常以对数形式 $10\lg(I_{max}/I_{min})$ 来表示. 大的调制深度对于长距离传输、低误码率及高接收敏感性非常重要.

3. 插入损耗

这一指标是指把调制器作为无源器件引入系统时，器件本身对输入光能量的损耗. 通常用输入和输出光强的比值来衡量. 显然调制器插入损耗应当被累积进系统的整体损耗中.

4. 比特能耗

比特能耗指的是传输 1 bit 的信息所需的平均能量. 它是衡量器件本身竞争力的一项重要指标. 比如说,如果要用硅基集成光互连系统来代替 32 nm CMOS 技术中的铜互连系统,光学系统的能耗必须不高于 100 fJ/bit. 这意味着对光学调制器这样的光输出器件,其能耗必须不超过 10 fJ.

电吸收型调制器可能提供更低的比特能耗. 因为它利用强量子限制斯塔克效应,不依赖于载流子浓度的变化,因而功耗相对较低. 因此 QCSE 型调制器是未来降低功耗的新调制器重要考虑之一.

5. 调制器几何尺度

这一指标考查调制器所耗用材料的多少,集成制造成本、功耗以及性能. 一般来说,调制器应当做得尽可能小,但不足以让调制机制失效.

6. 光学带宽

这是指调制器能正常工作的光波长范围. 不同的调制器结构设计,对波长变化的敏感度不同. 例如,MZI 型调制器工作波长范围比共振结构型调制器宽.

7. 温度和工艺的敏感度

这一指标对不同设计结构的调制器差别很大. 我们尽量追求对制造工艺要求不那么高,在 $-5\sim70$℃ 范围内稳定工作的集成光调制器.

显然在以上指标中,最重要的是调制速度/调制带宽和几何尺度. 这是截至目前研究最着重的. 但随着调制速度逐渐接近可接受的水平,功耗等其他指标也会变得越来越重要.

§6.3 硅基电光调制

一般来说,当外加电场作用在材料上时,其折射率的实部和虚部都会发生变化. 我们将实部的变化 Δn 称为电致折射,而将虚部的变化 $\Delta\alpha$ 称为电致吸收. 前面提到的泡克耳斯效应和克尔效应均属于电致折射的内容. 由于硅材料具有中心对称的晶体结构, 没有泡克耳斯效应,而且克尔效应也非常微弱,因此,通过电致折射的方法来进行电光调制是不实际地. 如果希望利用硅来进行电光调制,必须采用别的方法来更有效地改变硅的折射率. 本节将阐述几种有效改变硅的折射率的方法,也就是电致吸收的方法,以及相对应的电光调制机理,并介绍两种最常用的硅基电光调制器.

6.3.1 硅的光吸收及电光效应

像其他材料一样,硅材料也会对通过它的光进行吸收并产生损耗. 其主要的吸收方式有两种: 一种是带边吸收,发生在入射光波长不超过由导带和价带带

隙决定的截止波长的情况下. 一般来说,波长越短,吸收越强,如图 1.5 所示. 这样的光被吸收后可以激发电子从价带到导带的跃迁,是个强吸收过程. 另一种称为自由载流子吸收,是材料中的自由载流子对输入光的吸收. 改变自由载流子浓度,将会改变对光的吸收,从而影响到折射率的实部和虚部. 在入射光波长长于截止波长时,这一过程将占主导地位. 下面介绍几种常见的与电致吸收有关的电光效应.

1. 弗朗兹-凯尔迪什(Franz-Keldysh,F-K)效应

当存在外电场时,某些半导体材料的能带发生弯曲,导致导带和价带间的带隙发生变化,从而影响晶体的光吸收特性,特别是对那些波长与带隙宽度相当的光. 这种影响不仅体现在电致折射上,也体现在电致吸收上,但效果都不强烈. 图 6.1 显示了室温下硅材料在 1.07 和 1.09 μm 波长附近的电致折射特性(电致吸收的关系图)[3]. 可以看见,即使在高达每厘米上万伏的电场强度下,光学特性的改变仍然很小. 值得注意的是,这种效应随着波长的增长而减小. 在 1.31～1.55 μm 的通信波长范围内,光学特性的变化几乎可以忽略不计.

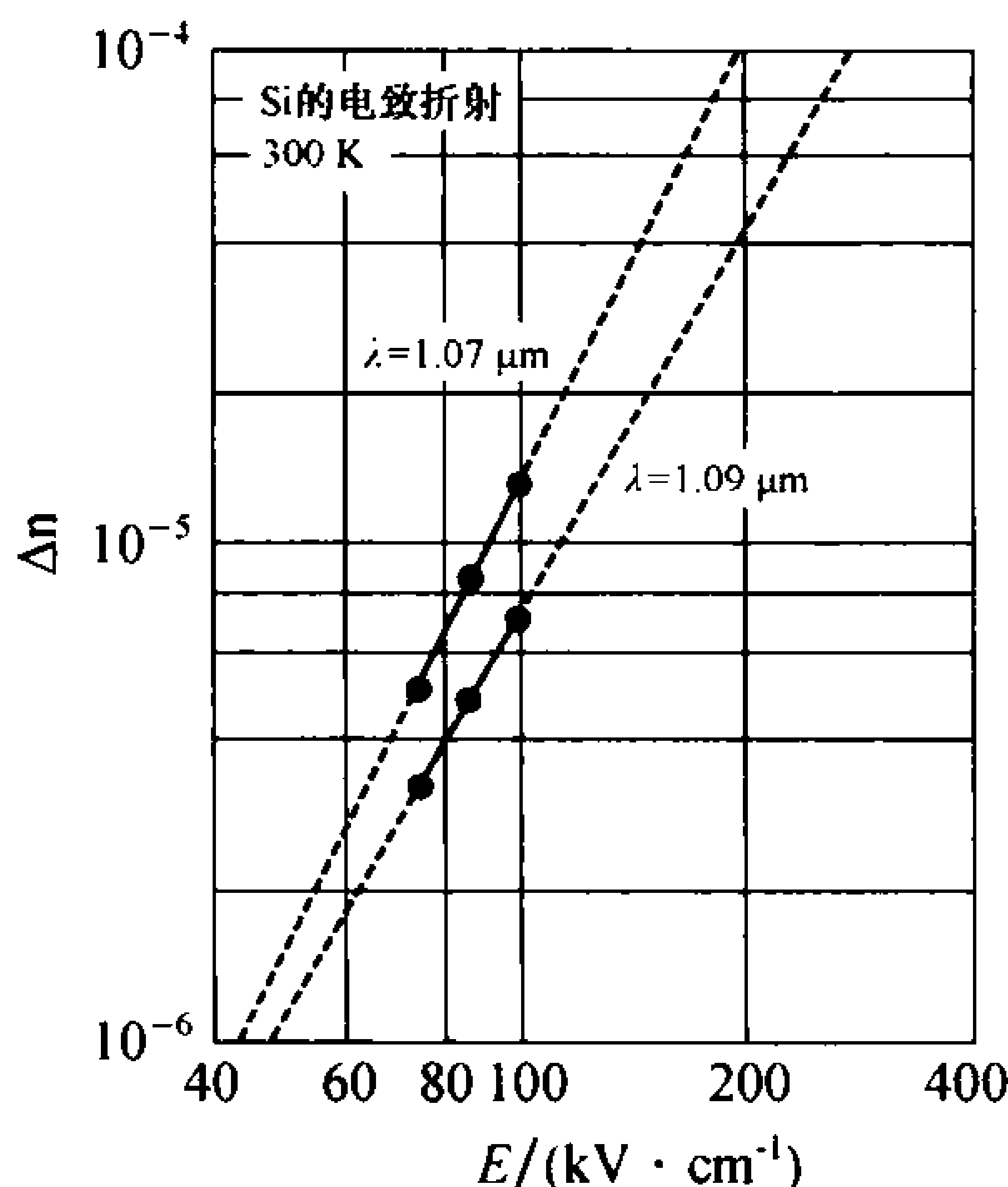


图 6.1　电致吸收关系图

2. 量子限制斯塔克效应(quantum confined stark effect,QCSE)

量子限制斯塔克效应反映了量子阱结构光学吸收谱在外加垂直电场作用下的变化. 在没有外加电场的情况下,量子阱结构中的电子和空穴只占据了子能带结构中的部分离散能级,因此,该量子阱结构只能吸收相对应的一系列离散光波. 然而,当加上外加电场时,电子能级向下移动,空穴能级向上移动,从而使该结构能够吸收的波长向长波长方向扩展. 显然量子限制斯塔克效应也是通过电场的存在改变材料对光的吸收特性. 由于都是通过外加电场来改变材料的光学

吸收特性，QCSE 和 F-K 效应具有内在的一致性. 由于狭窄量子阱的存在，其势垒将允许激子存在并增强光的吸收(QCSE). 当量子阱宽度增大时，这种增强效应将迅速减弱，其极限状态就是 FK 效应.

3. 等离子色散(plasma dispersion, PD)效应

理论与实践已经证明，材料中自由载流子的浓度对其折射率的实部和虚部都会产生影响，进而影响材料的光学性能. 自由载流子的浓度对材料折射率虚部(吸收)的影响远大于对实部的影响. 其对虚部(吸收)的影响可以利用人们熟知的德鲁德-洛伦兹(Drude-Lorenz)公式来描述：

$$\Delta\alpha = \frac{e^3\lambda_0^2}{4\pi^2 c^3 \varepsilon_0 n}\left[\frac{\Delta N_e}{\mu_e (m_{ce}^*)^2} + \frac{\Delta N_h}{\mu_h (m_{ch}^*)^2}\right], \tag{6.3.1}$$

其中 $\Delta\alpha$ 为材料折射率中吸收系数的变化量；e 为电子电量；λ，c，ε_0 分别为真空中的光波长、光速和介电常数；n 为本征折射率；ΔN_e 和 ΔN_h 分别代表材料中的自由电子和空穴浓度；m_{ce}^* 和 m_{ch}^* 分别为电子和空穴的有效质量；μ_e 和 μ_h 分别为电子和空穴的迁移率.

通过 K-K 关系式

$$n_r(w) - 1 = \frac{c}{\pi} P\int_0^\infty \frac{\alpha(w_1)}{w_1^2 - w^2} \mathrm{d}w_1, \tag{6.3.2a}$$

$$\Delta n_r(w) - 1 = \frac{c}{\pi} P\int_0^\infty \frac{\Delta\alpha(w_1)}{w_1^2 - w^2} \mathrm{d}w_1, \tag{6.3.2b}$$

可以将自由载流子的浓度对材料折射率实部的影响直接表达为下式：

$$\Delta n = \frac{-e^2\lambda_0^2}{8\pi^2 c^3 \varepsilon_0 n}\left(\frac{\Delta N_e}{m_{ce}^*} + \frac{\Delta N_h}{m_{ch}^*}\right). \tag{6.3.3}$$

公式(6.3.1)和(6.3.3)就是所谓的等离子色散效应. 它表达了材料中自由载流子的浓度变化所引起的光学特性的变化之间的定量关系. 如果通过电场的改变来注入或减少载流子浓度，就可以实现电光调制.

特别地，如果利用硅材料对通信波段的光波进行调制，大量的文献数据已被总结成以下的经验公式[4]：

在 1.55 μm 波段，有

$$\begin{cases} \Delta n = \Delta n_c + \Delta n_h = -8.8\times 10^{-22}\Delta N_c - 8.5\times 10^{-18}(\Delta N_h)^{0.8}, \\ \Delta\alpha = \Delta\alpha_c + \Delta\alpha_h = 8.5\times 10^{-18}\Delta N_c + 6.0\times 10^{-18}\Delta N_h. \end{cases} \tag{6.3.4}$$

在 1.3 μm 波段，则有

$$\begin{cases} \Delta n = \Delta n_c + \Delta n_h = -6.2\times 10^{-22}\Delta N_c - 6.0\times 10^{-18}(\Delta N_h)^{0.8}, \\ \Delta\alpha = \Delta\alpha_c + \Delta\alpha_h = 6.0\times 10^{-18}\Delta N_c + 4.0\times 10^{-18}\Delta N_h. \end{cases} \tag{6.3.5}$$

如果在 1.55 μm 波段设定 $\Delta N_e = \Delta N_h = 5\times10^{17}$，我们有

$$\Delta n = -8.8\times10^{-22}\times(5\times10^{17}) - 8.5\times10^{-18}\times(5\times10^{17})^{0.8}$$
$$= -1.66\times10^{-3}. \quad (6.3.6)$$

显然，电场对硅材料光学性质的影响可以总结如下：泡克耳斯效应是零，F-K 效应几乎是零，克尔效应是 10^{-4} 数量级，而等离子体色散效应则在 10^{-3} 数量级. 这就是为什么多数硅基电光调制器都是通过等离子体色散效应来实现的.

6.3.2 硅基电光调制机理

本节将首先从硅光学调制所利用的物理效应出发，阐释目前常用的几种调制过程. 其次，具体分析了几种典型调制器的设计结构和优缺点，从作用原理和机制上给以解释，并特别针对等离子体色散效应型调制器给出详细的分析，揭示耗尽型结构为什么能获得目前的最高调制带宽的原因. 最后，对其他新型调制机理和调制器设计做简要介绍.

1. 硅基电光调制机制

硅材料中的电光效应包括克尔效应、F-K 效应、QCSE 和 PD 效应等. 由于体硅材料中克尔效应和 F-K 效应都非常微弱，因此，硅基高速电光调制一般都利用硅材料的 PD 效应[3]. 等离子色散效应是一种间接电光效应，它利用外加电场下有源区自由载流子浓度的变化来调制输出光波的幅值和相位从而实现电光调制. 但是，由于受到载流子本身的复合寿命的限制，器件的开关速度只能达到兆赫量级[5]. 要想提高硅基电光调制的速度，必须寻求新的调制器结构.

光学调制是对传播光束的基本特征进行调制. 根据被调制特征的不同，可以简单分成幅度，相位和偏振调制. 另一种划分方法是分成电致折射调制和电致吸收调制. 施加外电场导致材料复折射率的变化. 其中导致实部变化 Δn 叫做电致折射，导致虚部变化 $\Delta\alpha$ 叫做电致吸收. 通常我们就是利用前面提及的克尔效应、F-K 效应来进行电光折射调制或者实现电光吸收调制. 但在硅中，这些效应都非常微弱. 我们只能使用等离子色散调制.

对于直接电致吸收调制，这过程非常简单，只要改变通过电控吸收与否，改变出射光强度就可以实现. 对于电致折射，我们有两种间接方法仍旧可以实现光的强度调制. 第一种方法是折射率改变可以用来调节两路传播光的相对相位，从而使得它们实现相长或者相消干涉. 这就实现了光强度调制. 典型地，一个 M-Z 干涉仪就可以用来实现这种调制. 第二种实现把折射率变化转化成光强度变化的方法是引入一个共振环结构，调制器折射率变化必然改变共振环与光传播路径的共振条件，由此实现设备在共振与非共振间的切换.

这样，我们就把电场在材料中导致的光性质的变化都最终能转化为光强的按电场的变化. 这是电光调制的最基本过程. 具体一个调制器的实现必然是按照基

本物理过程和物理原理，并同时应用我们上述的两种机制之一来工作的. 常见的硅集成高速调制用的是等离子色散物理效应，调制机制用的是电致折射调制.

2. 典型调制机制分析

早期硅波导光调制器都是利用等离子色散效应，利用载流子注入来改变波导材料折射率或者吸收系数. 环绕波导形成的 p-i-n（p-type/intrinsic/n-type layers）二极管结构用来控制光传播路径中的空穴和电子注入. GHz 带宽的方案在 20 世纪中期就被提出. 图 6.2 是 Png 等提出的一个方案[8].

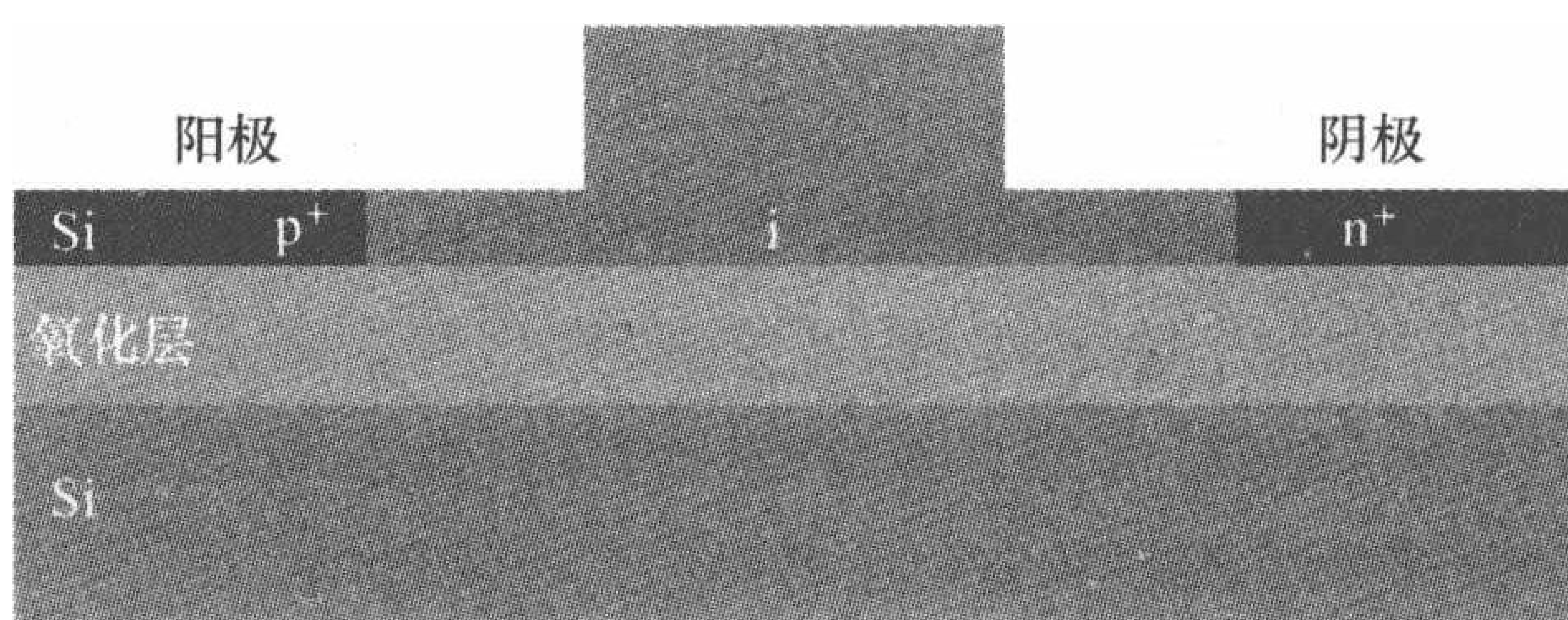


图 6.2 载流子注入型调制器示意图

突破 1 Gbit 调制速度也是在 2004 年实现的，见图 6.3. 该结构与以往不同的是波导中的载流子浓度不再是通过电控注入来改变，而是通过电控来积累. 在微米尺度波导里面构造了一个介层，自由载流子在介电层两侧积累（这非常像电容器两极板上电荷的积累）. 这一载流子积累型调制器优化后调制速度达到 10 Gbit/s. 最近 Lightwire 公司演示了它的设备在 10 Gbit/s 调制速度下，消光比仍能达到 9 dB.[6]载流子积累型调制器调制速度不再受材料最小载流子寿命限制. 而是依赖于器件的电阻和电容. 这也是载流子积累型调制优于载流子注入型之处.

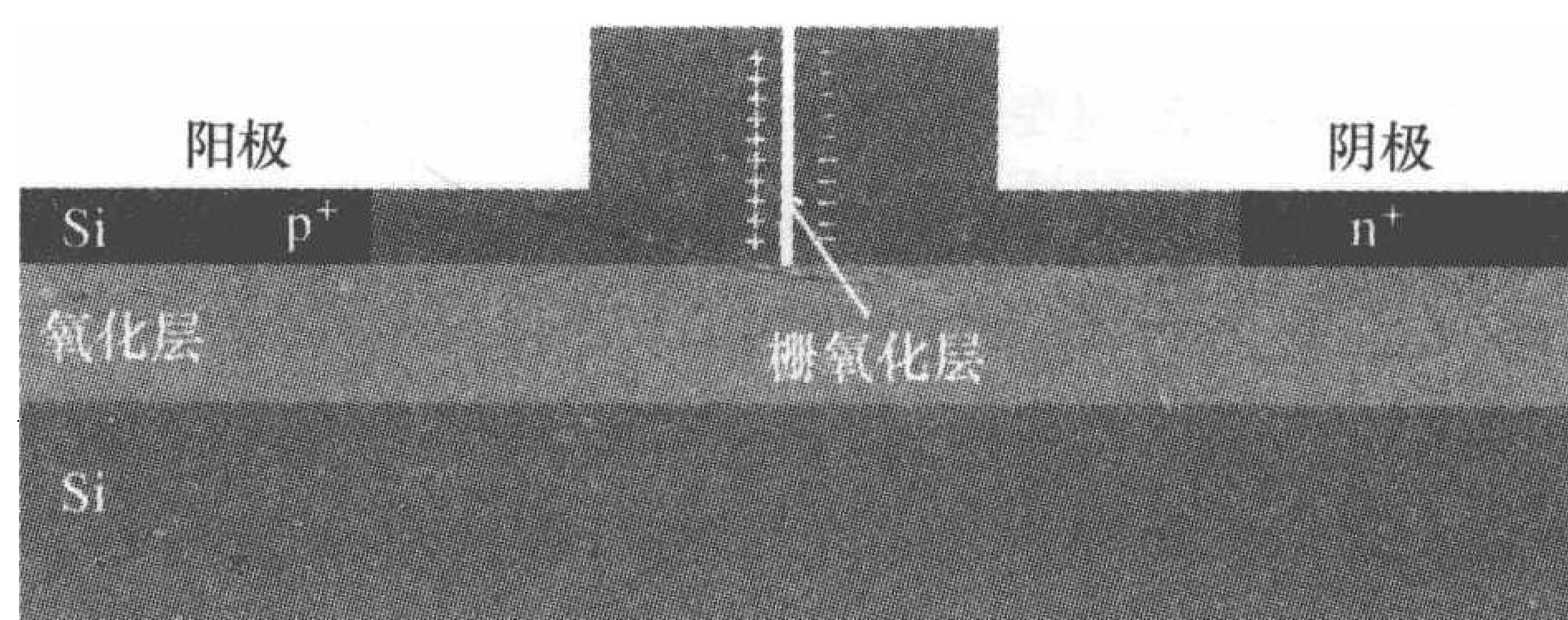


图 6.3 载流子积累型调制器示意图

后来出现的载流子耗尽方案也同样避免了载流子最小寿命的影响，如图 6.4 所示. 这一机制利用 p-n 二极管的反向偏压形成的耗尽层工作. 施加不同的反向偏压，形成不同宽度的耗尽区，从而光波导中的载流子密度自然不同. 2007 年，Intel 公司利用这种机制，首次实现了 30 Gbit/s 的光调制器. 此后继续实现了 40 Gbit/s 的调制.[9,15]

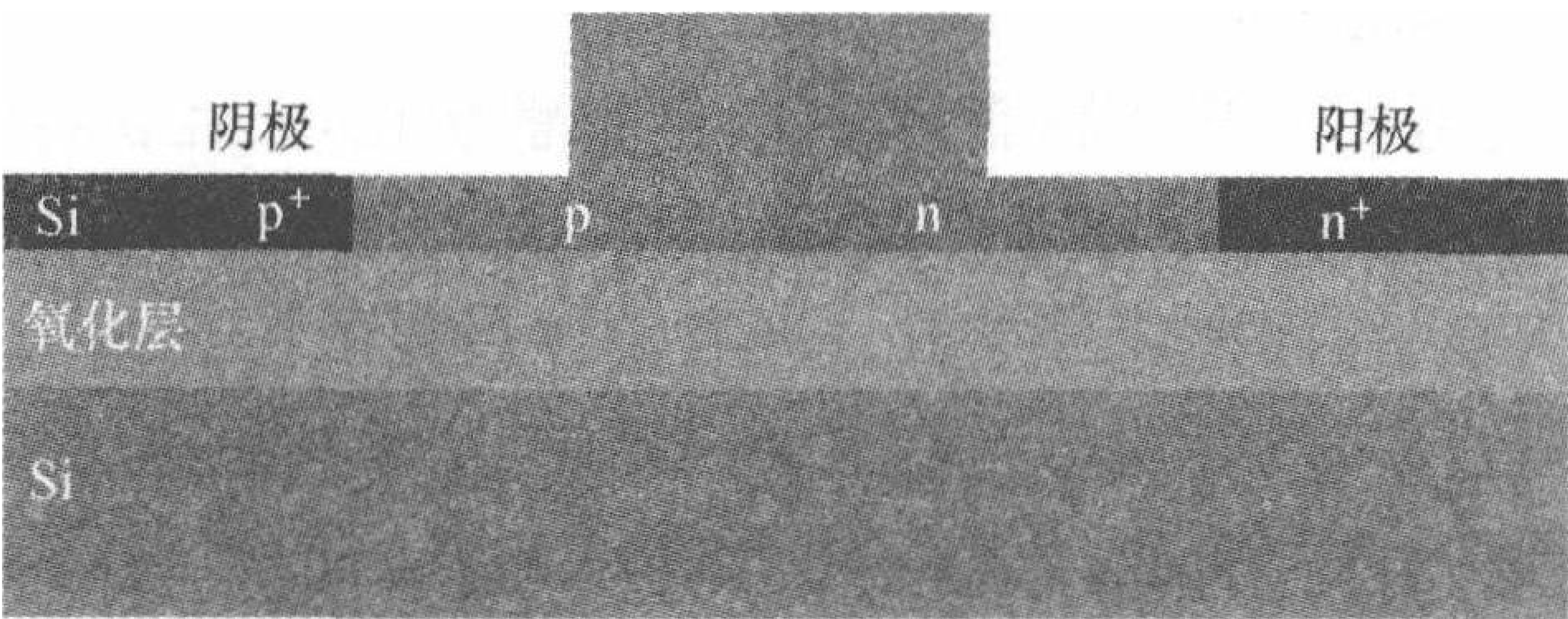


图 6.4　载流子耗尽型调制器示意图

通过以上分析可以得出：如果采用等离子色散机制，通过改变载流子浓度实现电光调制，那么提高载流子浓度变化速度是提高调制速率的关键. 以上三种结构(载流子注入，载流子积累，载流子耗尽)分别对应不同的微观载流子浓度变化机制，见表 6.1. 在目前较为常用的 MZI 型和共振型调制器中，大都采用以上三类调制机制，且以载流子耗尽型所获得的调制速度最高，见表 6.2[6].

表 6.1　硅基电光调制器典型调制机制比较

结构	特点
载流子注入	高掺杂的 p 区和 n 区被本征区分开. 波导在本征区；正偏压促使电子和空穴注入到本征波导区
载流子积累	薄 SiO_2 绝缘层把波导分成两半，形成电容结构；载流子的积累随所加电压的变化而变化
载流子耗尽	轻掺杂 p 区和 n 区相邻，形成 pn 结二极管；当施加的反向偏压增大时，二极管的耗尽区变大

表 6.2　目前几种调制器在不同调制机制下的性能比较

调制规律	结构	器件尺寸/μm^2	达到的速率/($Gbit \cdot s^{-1}$)	比特能耗/($fJ \cdot bit^{-1}$)	调制电压/V	直流调制深度和插入损耗/dB	调制深度与速率之比/($dB \cdot Gbit^{-1} \cdot s$)	工作光谱/nm
水平 p-n 结耗尽层	MZI	-10^4	40[39]	3×10^4	6.5	>20，−7	1/30 1/40	>20
正向偏置二极管	MZI	-10^3	10	-5×10^3	7.6	6~10，12	—	—
正向偏置二极管	Ring	-10^2	>12.5	−300	3.5	>10，<0.5	8/16 3/18	−0.1
反向偏置二极管	Disk	20	10	85	3.5	8，1.5	—	−0.1
正偏 p-i-n 二极管	Ring	-10^2	3	86	0.5	7，1	—	0.2
反偏二极管	Ring	−10	10	50	2	6.5，2	8/10	−0.1

3. 新型硅基调制机制

毋庸置疑，现有的硅基调制器件与10年以前相比，已经获得非常大的进步．但调制器作为一个整体性功能器件，目前其各种指标以及综合性能还远不能达到人们预期商用水平．因此在继续改进现有设计过程中，人们仍在试图发现新的调制机制和物理过程，从而获得性能上的突破．

在这一过程中，人们首先考虑和CMOS工艺兼容的材料，比如Ge元素．随之利用QCSE进行电吸收调制的调制器原型的演示．2005年，Kuo等展示了纯Ge量子阱和富Ge的SiGe势垒的QCSE[7]．尽管目前还没有波导型QCSE调制器演示，但高调制速度，高效率的QCSE调制器正在成为目前的研究热点，如图6.5所示[8]．其次，将来的调制也可能通过操纵硅晶体的结构．比如给晶体硅施加应力，使之丧失对称性，从而，使在对称晶体硅结构中无法存在的泡克耳斯效应得已体现．

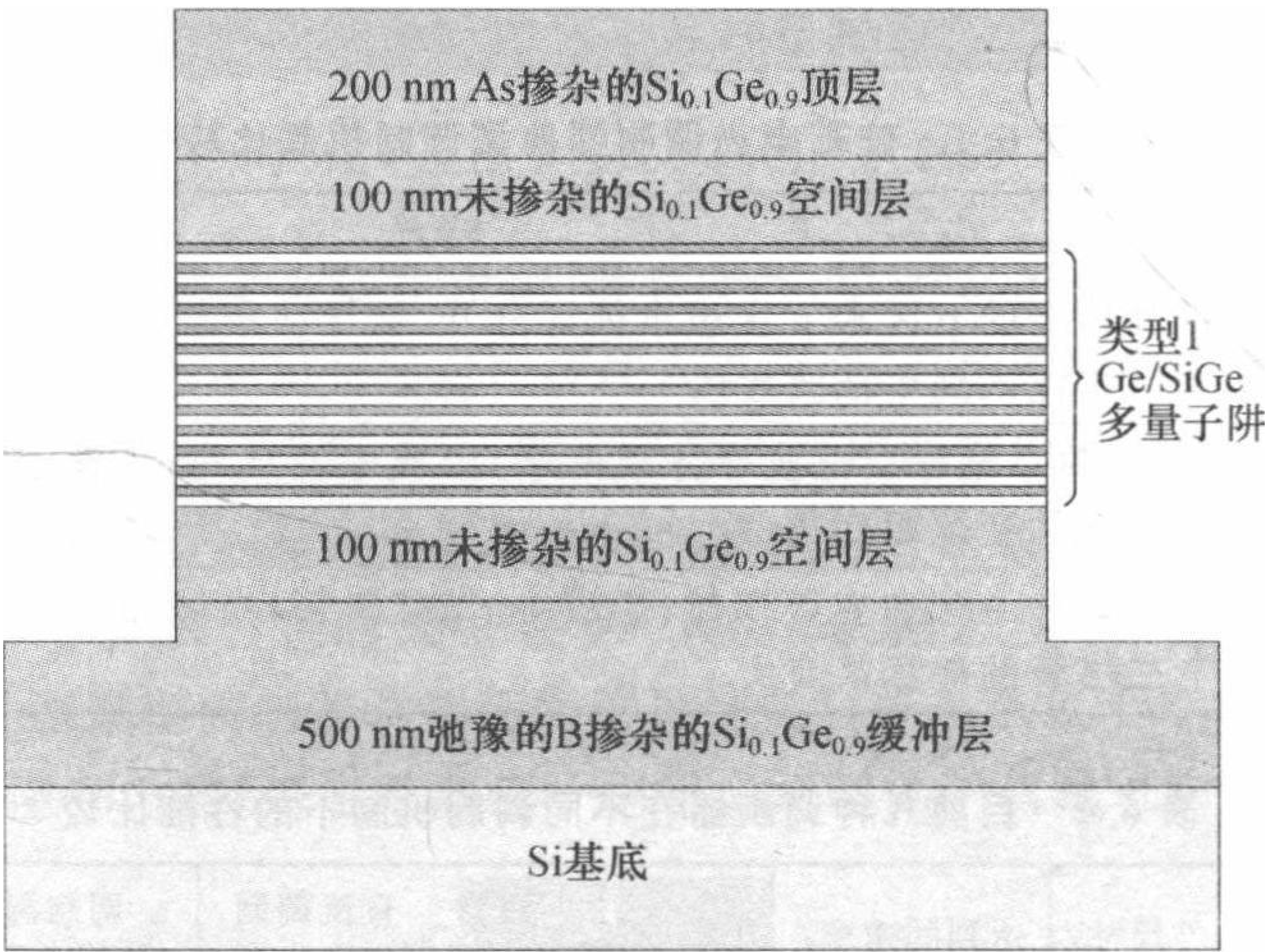


图6.5 基于量子限制斯塔克效应的锗硅调制器结构图

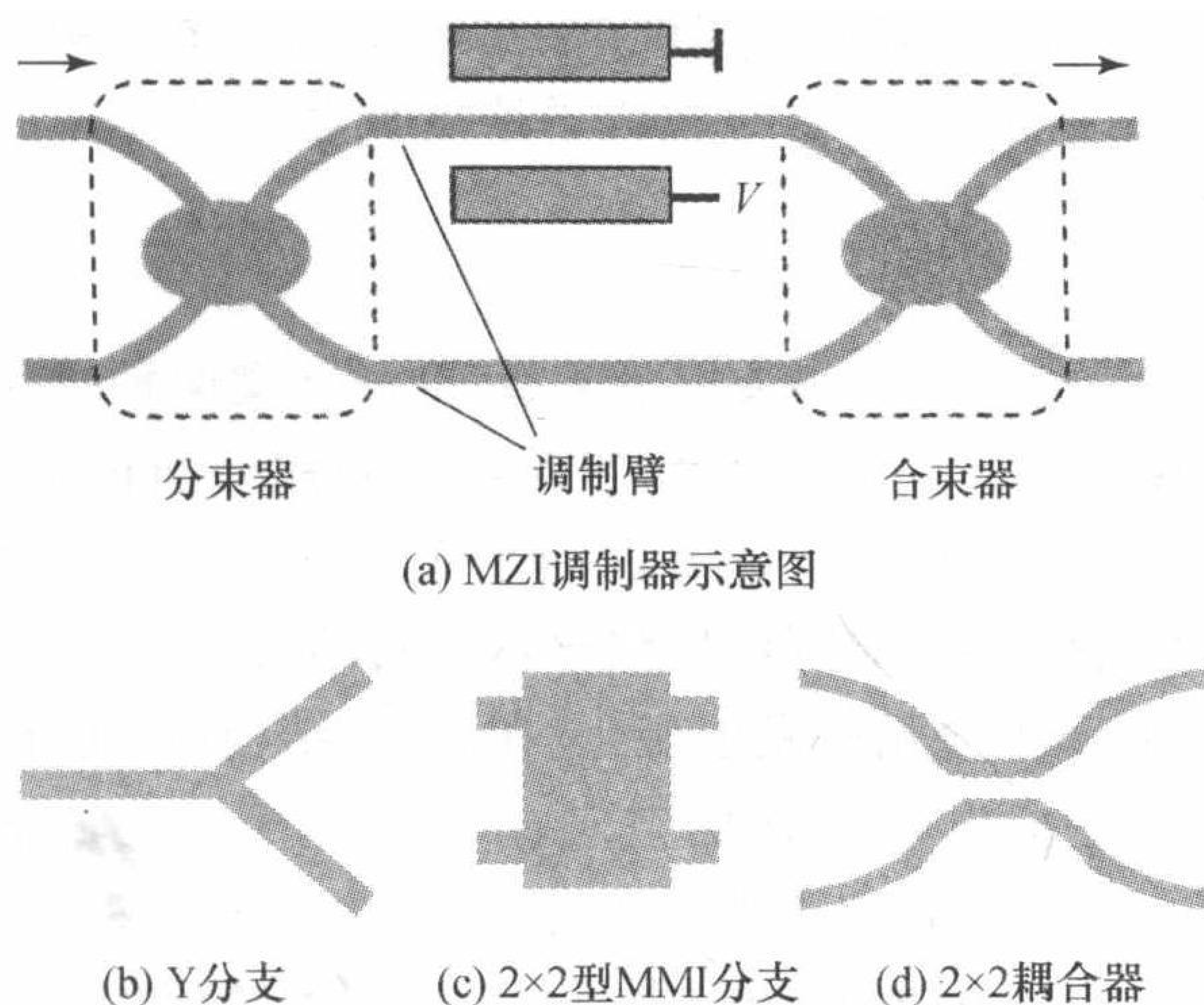
6.3.3 硅基MZI电光调制器

硅基电光调制器按电学结构划分主要包括p-i-n结构、三端结构和MOS电容结构等；按光学结构划分，有MZI结构、F-P结构、布拉格光栅结构和微环共振结构等[4]．后三种光学结构是基于输入波长谐振原理，而MZI结构则是基于光场干涉原理．相对于谐振结构，干涉结构的调制器具有较宽的调制带宽和较低的温度敏感性，因而具有更广泛的应用前景[6]．硅基MZI电光调制器是实现未来大容量、高数据密度传输的光网络和光互连通信的关键光电子器件，它的运用可以降低光网络和光互连系统的制作成本并且大大有利于未来的硅基光电集

成.目前国际上已经实现的硅基 MZI 高速电光调制器件的调制速率已经超过 40 Gbit/s.本小节对硅基 MZI 高速电光调制器的几种常见结构、原理、性能参数进行了介绍,对不同载流子调制方式的优缺点进行了比较和评述,然后总结国内外的最新研究进展.

1. 硅基 MZI 电光调制器的原理

M-Z 干涉型(MZI)调制器是利用相位调制实现强度调制的器件,广泛地应用于各种光调制器和光开关中,如铌酸锂调制器、硅基调制器等.MZI 结构通常由输入端、输出端、两条调制臂、两个分束器构成.分束器通常为 Y 分支、MMI 分支、或者 3 dB 耦合器.两调制臂一般为对称结构,也有非对称的情况.图 6.6(a)所示为常见的 MZI 波导调制器(modulation waveguides, MW),结构中的两条调制臂采用了载流子注入式结构、耗尽式结构或者 MOS 电容式结构的硅波导,图 6.6(b),(c),(d)为前面所述三种分束/合束器.以 Y 分支结构 MZI 调制器为例,当入射光从输入端输入,经过一个 Y 分支时,就会被分成两束,这两束强度相位都相同的光分别进入两条调制臂,两条调制臂是对称的,当没有外加电压时,这两束光的相位是完全一致的,在输出处的 Y 分支汇合时没有相位差,根据光波的干涉理论,当两束光的相位差为零时,干涉相长,从输出端输出光强最大,此时的光信号可以看做是“1”信号;当对其中的一条调制臂外加电压进行调制时,该调制臂的有效折射率发生变化,两束光的相位不一致,在输出端 Y 分支处汇合时,存在相位差,进行干涉后,光强不是最大值,若相位差为 π,干涉相消,在输出端没有光输出,那么此时可以看做输出“0”信号.



(a) MZI调制器示意图

(b) Y分支　(c) 2×2型MMI分支　(d) 2×2耦合器

图 6.6　M-Z 波导调制器结构示意图及三种结构的分束/合束器

下面以 MMI＋MW＋MMI 结构为例推导其数学表达式，说明其工作原理.

MMI 的传输矩阵为

$$\boldsymbol{S}_{\mathrm{MMI}}=\frac{1}{\sqrt{2}}\begin{bmatrix}\exp(\mathrm{i}\varphi_{11}) & \exp(\mathrm{i}\varphi_{21})\\ \exp(\mathrm{i}\varphi_{12}) & \exp(\mathrm{i}\varphi_{22})\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & \mathrm{i}\\ \mathrm{i} & 1\end{bmatrix}, \tag{6.3.7}$$

MW 传输矩阵为

$$\boldsymbol{S}_{\mathrm{MW}}=\begin{bmatrix}\exp(-\mathrm{i}\Delta\Phi/2) & 0\\ 0 & \exp(\mathrm{i}\Delta\Phi/2)\end{bmatrix}, \tag{6.3.8}$$

故 MZI 调制器的传输矩阵为

$$\boldsymbol{S}=\boldsymbol{S}_{\mathrm{MMI}}\boldsymbol{S}_{\mathrm{MW}}\boldsymbol{S}_{\mathrm{MMI}}=\begin{bmatrix}\sin(\Delta\Phi/2) & -\cos(\Delta\Phi/2)\\ -\cos(\Delta\Phi/2) & -\sin(\Delta\Phi/2)\end{bmatrix}. \tag{6.3.9}$$

在外加调制时，光通过调制臂后，产生的相位差为 $\Delta\Phi=\Delta\beta L$，其中 L 为调制臂的长度，传输常数差 $\Delta\beta=k_0\Delta n_{\mathrm{eff}}=2\pi\Delta n_{\mathrm{eff}}/\lambda$，由此可以得出 MZI 结构调制器中，波导层的有效折射率改变在两调制臂上产生的相位差

$$\Delta\Phi=2\pi\Delta n_{\mathrm{eff}}L/\lambda. \tag{6.3.10}$$

设输入 MMI 上端的输入端光强为 I_{in}，下端输入端光强为 0. 则选取输出 MMI 上端输出端作为调制器输出波导，其光强为

$$I_{\mathrm{out}}=I_0\sin^2(\Delta\Phi/2)=I_0\sin^2(\pi\Delta n_{\mathrm{eff}}L/\lambda). \tag{6.3.11}$$

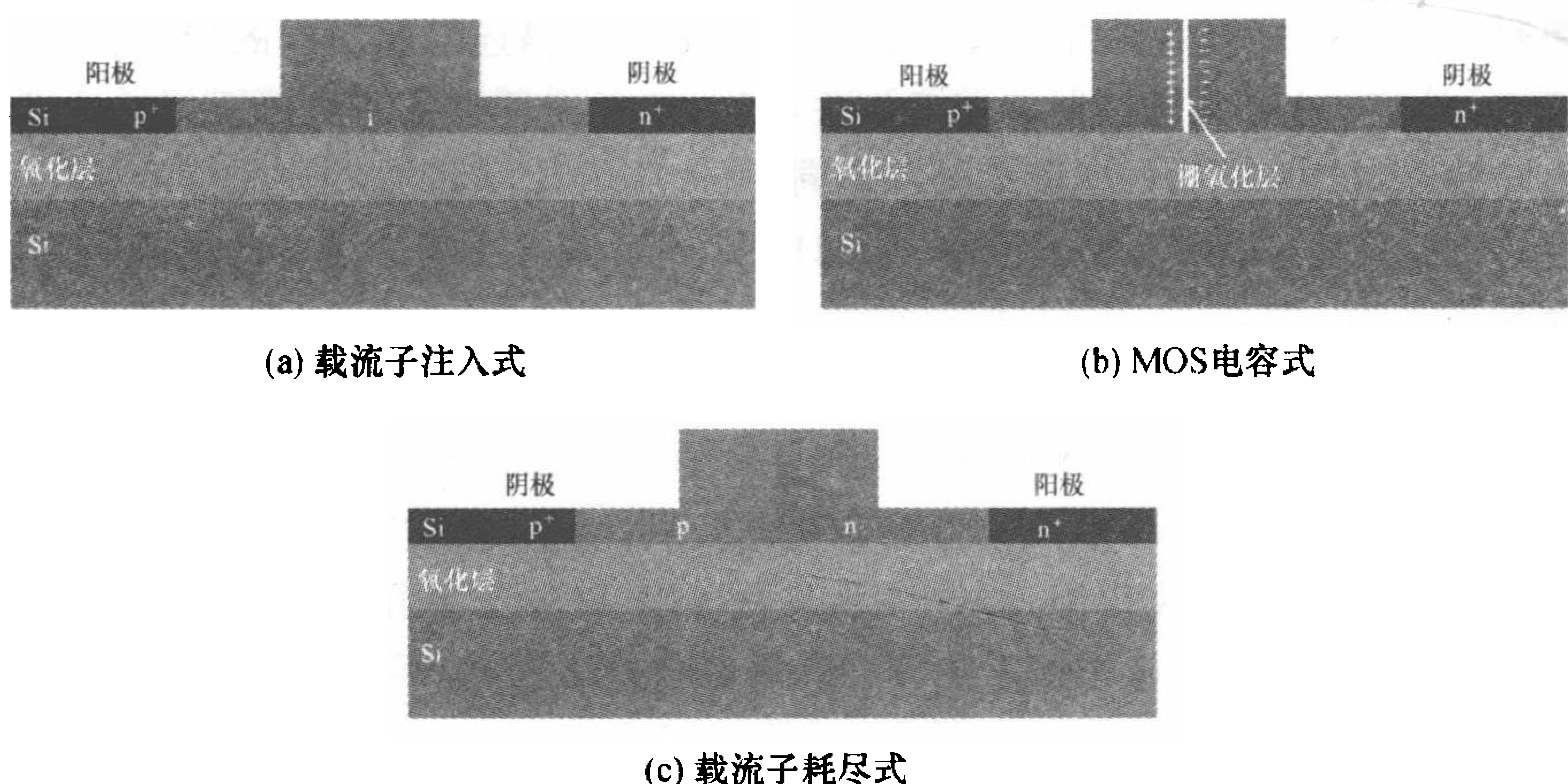
改变调制电压，即改变 Δn_{eff}，I_{out} 也发生改变，从而实现了光强度的调制. 通过合理的设计调制区结构，可对调制器的调制性能进行优化.

2. 硅基电光调制器的电学结构

在硅基电光调制中，对载流子浓度调制的方式有三种：载流子注入式、载流子积累式、载流子耗尽式[6]. 根据对载流子浓度调制方式的不同，可以将硅基电光调制器分别归入这三类.

(1) 载流子注入式.

载流子注入式的电光调制主要有 p-n 型和 p-i-n 型两种. 在 p-n 结构中，p 区和 n 区的交界处存在一个空间电荷区，当对器件施加外加正向电压时，会在 p-n 区之间建立电场，在该电场作用下，p 区的空穴向 n 区移动，n 区的电子向 p 区移动，使得中间的空间电荷区减薄，导致该区域的载流子浓度发生变化，根据等离子色散效应，折射率也会相应发生变化. 在 p-i-n 结构中，在外加正向偏压的情况下，在 p-n 区之间建立电场，p 区的空穴、n 区的电子在电场作用下进行迁移，注入中间的 i 区，使得 i 区载流子浓度发生变化，引起波导折射率发生变化. 图 6.7(a)给出了硅基 p-i-n 型电光调制器示意图. 这类结构的调制器折射率改变大，调制深度深，但受制于载流子的碰撞、复合、迁移和载流子寿命的因素，因此调制速度不高.



(a) 载流子注入式　(b) MOS电容式

(c) 载流子耗尽式

图 6.7　硅基电光调制器的三种基本结构

(2) 载流子积累式.

图 6.7(b)是载流子积累式硅基电光调制器的典型结构，它的核心部分是由一个栅介质层和分别居于栅介质层两侧的 p 型掺杂波导和 n 型掺杂波导构成的 MOS 电容结构，外加正向偏压时，由于电容的存在，在栅介质层一侧的 p 型 Si 波导区会有空穴聚集在栅介质层附近，另一侧的 n 型 Si 波导区会有电子聚集在栅介质层附近，即在栅介质层附近的波导会发生载流子浓度的变化，导致硅波导的折射率改变，从而实现调制. 由于载流子浓度发生变化是由感应而来的，不存在载流子的漂移过程，因此 MOS 电容式的电光调制器速度调制很快，但是由于其载流子变化的区域非常小，和光场重叠的区域就非常小，因此器件的调制效率低.

(3) 载流子耗尽式.

载流子耗尽式的电光调制器是利用 p-n 结的空间耗尽区随外加反向偏压的大小而变化，从而改变载流子的浓度，实现对折射率的调制[9]. 图 6.7(c)为反向 p-n 结结构硅基电光调制器示意图. p-n 结区存在一个很薄的耗尽层，当外加反向电压时，结区的电子空穴分别在电场作用下向正负极运动，导致耗尽区厚度增大，结区载流子浓度发生改变，折射率改变. 此种结构的硅基调制器载流子寿命短，调制速度快，能达到几十 GHz，但是由于载流子浓度变化较小，变化区域较小，因此调制效率不高.

以上几种载流子的改变方式各有优缺点，可以根据具体应用需求进行选择.

3. 硅基 MZI 电光调制器的主要性能参数

(1) 调制带宽或调制速度.

调制带宽或调制速度是调制器的最重要的性能指标之一. 调制带宽定义为

调制效率达到50%的调制信号频率范围.调制速度描述了调制器的传输数据的速率.调制速度通常和调制带宽正相关,宽的带宽将具有更高的数据传输速率.不同的应用对调制器的速度的要求不同,某些应用,如光互连和光通信需要高密度的数据传输,因而对调制器的调制速度要求高,而光学传感则对调制速度要求较低.

(2) $V_\pi L_\pi$ 值.

$V_\pi L_\pi$ 值是评价MZI型硅基电光调制器性能的一个重要指标,它表征了器件的调制效率.其中 V_π 是调制器的半波电压,L_π 是调制器实现 π 相位变化的所需要的长度.在实际应用中,器件尺寸越小越好,以便于更好的集成化和更低插入损耗和功耗.$V_\pi L_\pi$ 越小即意味着器件的调制效率越高,所需的器件尺寸越小.

(3) 插入损耗.

插入损耗定义为增加调制器给光路带来的额外损耗.它是一种无源损耗,包括反射、吸收和模式耦合损耗,对光传输链路功率预算很重要.硅基光子学的一个重要方向是减小波导尺寸到亚毫米量级,然而小的横截面积却可能导致耦合损耗的增大.而且,由于载流子的存在,就会发生自由载流子吸收产生损耗.因此在对电光调制器进行设计时,要考虑到载流子吸收带来的损耗,选择合适的结构、掺杂浓度和掺杂区域.

(4) 功耗.

对于任何有源器件来说,功耗都是一个重要的性能指标,功耗低才能适应器件集成化的需要.这里功耗定义为传输每比特数据消耗的能量.光互连时代,由于高密度的数据传输,因而,功耗特性显得尤其重要.目前对于硅基电光调制器来说,功耗还不能做到很低,一般在几百个毫瓦量级.

4. 研究现状

下面分别按载流子调制方式的不同,对国内外高速电光调制器的研究现状进行总结.

(1) 载流子注入式结构MZI高速电光调制.

载流子注入式是硅基调制器最常见的结构,它通过给围绕波导的p-i-n或p-n结构加正向电压偏置,来控制电子和空穴向光传播区的注入.对p-i-n结构,为了避免额外的光场损耗,波导通常被限制在本征区,从而使泄漏到掺杂区的场模能量最小.早期,这种结构的调制器的带宽只能达到MHz量级,近几年,通过优化结构和减小器件尺寸,器件的带宽达到了GHz量级[10],消光比达到了10 dB[11].

(2) MOS电容结构M-Z高速电光调制.

硅基调制器研究的重大突破发生在2004年,Intel研究小组将MOS电容结构引入调制器的设计,首次实验实现了速率超过1 GHz的单片硅基调制

器[12]. 它的载流子调制方式不是采用载流子注入的原理，而是电荷感应，即微米尺寸的介质波导的两个表面发生电荷感应，形成类似于 MOS 电容的结构. 2005 年，Liao 研究小组对 MOS 电容结构的调制器进行了优化，使得其速度达到了 10 Gbit/s，消光比达到了 3.8 dB[13]. 之后基于 MOS 电容结构的硅基调制器被广泛报道，其中最引人关注的是 Lightwire 的报道，一方面，在速率达到 10 Gbit/s 的情况下，消光比达到了 9 dB[14]，另一方面，相对于 Liao 的结果[14]，他们的器件的调制效率提高了一个数量级，有源区长度从 3.45 mm 减小到了 480 μm. 与载流子注入式调制器不同，MOS 电容式调制器的速度不受载流子最短寿命的限制，而是决定于器件的电阻电容的大小.

（3）载流子耗尽式结构 MZ 高速电光调制.

载流子耗尽式结构是克服调制器速度限制的另一个重要技术，它也是通过降低载流子的寿命实现的[6]. 它的主要特点是给 p-n 结加反偏电压，光场和 p-n 结互作用，耗尽区宽度和载流子浓度随着反向偏压的变化而变化. 理论上，反偏 p-n 结的本征响应时间只有 7 ps[15]，但实验证明高频调制受限于调制器本身的电容和电极特性而难以提高[16]. 2005 年，英国萨里大学 Gardes 等首次提出了利用载流子耗尽式结构的调制器，如图 6.8(a)所示，并理论预测调制带宽可达 50 GHz[17]. 2007 年，Intel 公司 A. Liu 研究小组实现了这种结构的调制器[参见图 6.8(b)]，其 3 dB 带宽和数据传输率分别达到 30 Gbit/s[15]，不久又分别提高到了 30 GHz 和 40 Gbit/s[参见图 6.8(c)][10]，他的方法是结构上优化设计如设计行波电极和单边 n^{++} 区设计，以降低器件电阻电容. 之后，基于载流子耗尽结构实现速率大于 10 Gbit/s 的调制器被广泛报道[16,18]. 与载流子注入式结构调制器相反，这种结构的调制器具有高的调制速度（超过 30 GHz）和低的调制效率. 因此，存在的挑战是如何尽可能地提高其调制效率，同时保持高的调制速度等优点.

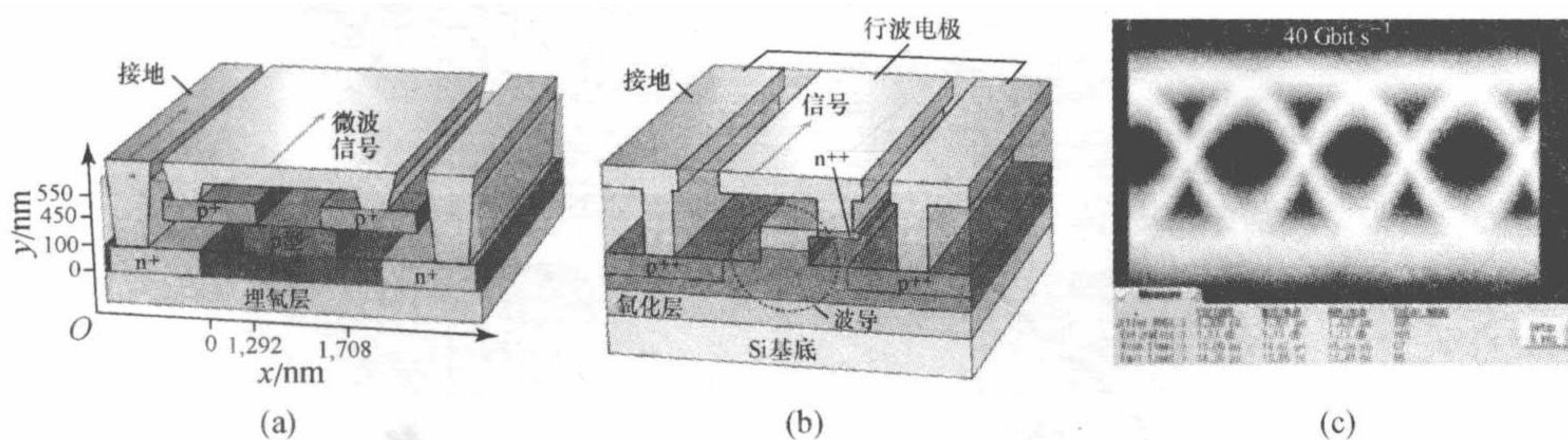


图 6.8　载流子耗尽式结构调制器结构和结果眼图

（4）光子晶体结构 M-Z 高速电光调制.

光子晶体的一个重要特性是慢光效应. 即光子在光子晶体中运动时，其群速

度 $d\omega/d\beta$ 会降低，直至趋于零．这样，如果人为破坏光子晶体的周期性结构，如在光子晶体中移去一些电介质，便可以产生缺陷．只有和缺陷态频率吻合的光子才能被局限在缺陷位置或只能沿缺陷位置传播，当偏离缺陷位置时，光就会迅速地衰减．这样一来，在光子晶体中就形成了点缺陷态或微腔．如果是线缺陷态，就可形成光波导．

给硅波导加电压，载流子浓度的变化导致硅波导的折射率变化 Δn，色散曲线将垂直移动 $\Delta\omega_0$．若硅波导采用光子晶体结构，由于群速度 $d\omega/d\beta_{pc}$ 显著降低，光子晶体波导的传播常数 $\Delta\beta=\Delta\omega d\beta_{pc}/d\omega$ 将显著增大．由相位的变化和传输常数及波导长度的关系 $\Delta\varphi=\Delta\beta_{pc}L$，光子晶体这样一种超常的 $\Delta\beta_{pc}$ 恰好可以极大的增强相位调制效率．可以很容易的使用光子晶体波导将 $\Delta\beta_{pc}$ 增大 100 倍，由此得到的结果是我们可以使用一个比传统波导短 100 倍的光子晶体波导来实现相同的相位改变．小尺寸是调制器追求的一个目标，长度只有十几微米的光调制可以降低器件的传输损耗和功耗．

运用光子晶体慢光效应的 M-Z 电光调制器最先由 Soljacic[19] 提出，之后基于载流子注入/耗尽和电荷感应等原理的调制器相继得到报道[20~22]，图 6.9[23] 所示是几种利用光子晶体波导的调制器结构，图 6.10 为文献[23]得到的实验结果．采用光子晶体结构的调制器的主要优点是能够通过控制光子群速度，从而使器件尺寸降到 100 μm 数量级，但是这通常需要以降低工作带宽为代价[24]．实际上，光子晶体的超常色散特性也可以用来改善器件工作带宽到 10 nm(1.24 THz)数量级，从而便可以在更宽的波长范围内获得温度不敏感性．但是，迄今报道的光子晶体结构的调制器的最大工作速度只有 1 GHz.

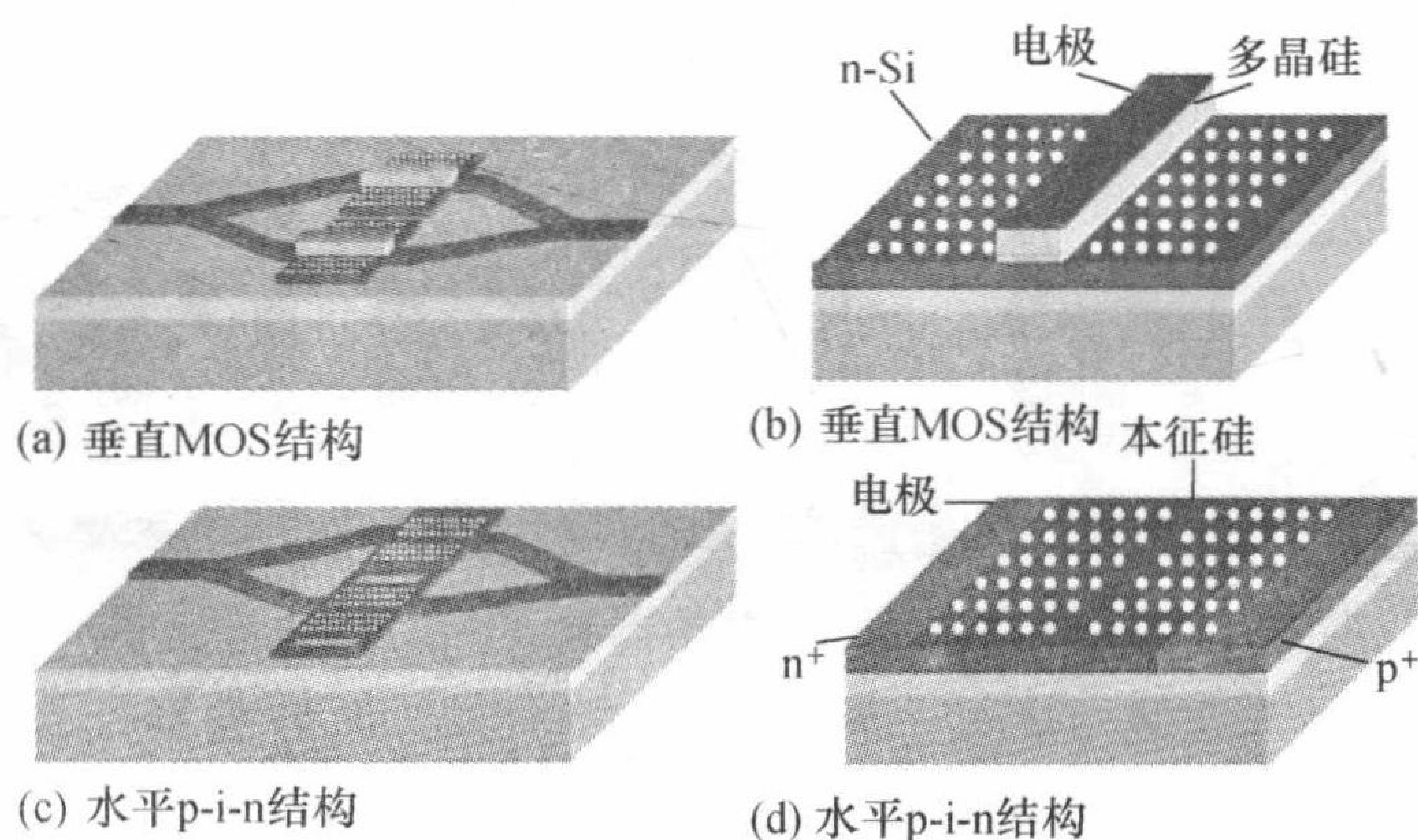


图 6.9　不同结构的光子晶体波导 MZ 调制器结构示意图

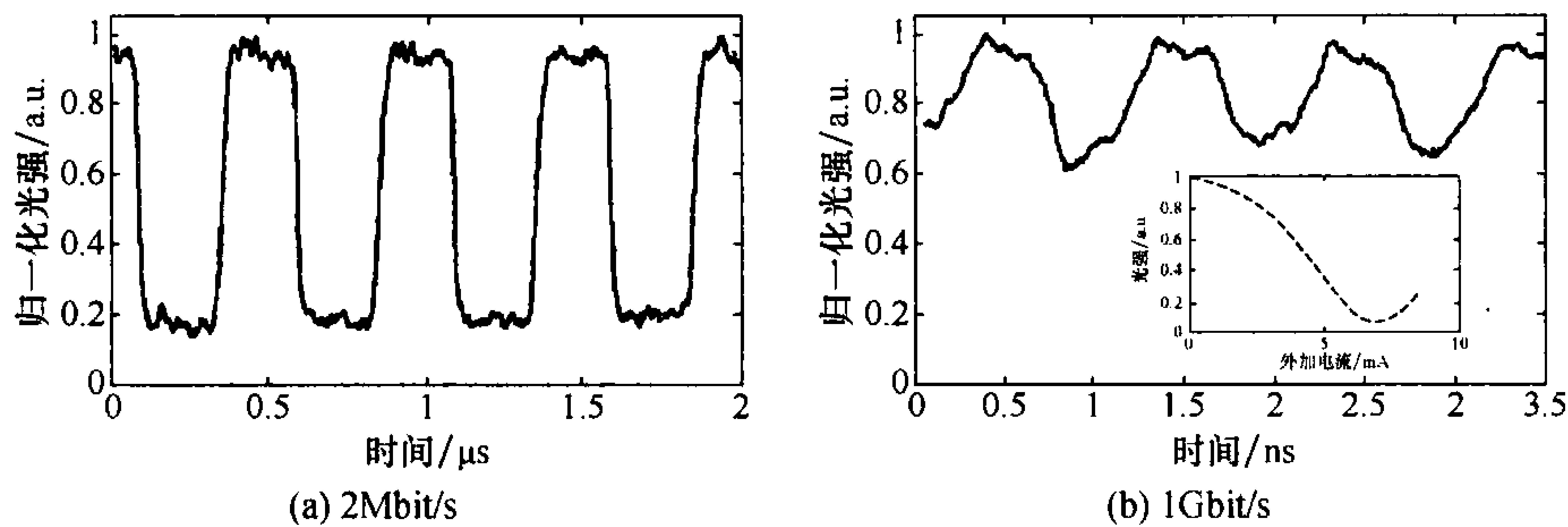


图 6.10　归一化光强测试结果

(5) 采用其他技术的硅基 MZ 电光调制.

除了对上面提到的基于三种载流子调制方式的调制器进行优化外，最近一些新颖的调制器得到了报道，他们或者运用新的电光效应，或者运用新的材料，或者运用新的结构，并获得了很好的性能. 文献[7,8] 报道了基于 QCSE 和 F-K 效应的 GeSi/Si 量子阱结构电光调制器. 文献[25]报道了基于克尔效应的电光调制器，它利用聚合物包层来克服硅里面的弱非线性效应，器件带宽达到了 10 GHz.

另外，前面的报道的电光调制器都存在调制带宽和调制效率之间的矛盾. 为克服这个矛盾，一些新型电光调制器相继被提出. 文献[26]报道的混合硅基电致吸收调制器实现了 10 dB 的消光比、30 nm 的光学带宽和 16 GHz 的调制带宽. 文献[27]报道了利用混合量子阱结构硅基调制器(参见图 6.11)，速率达到 10 Gbit/s，调制效率达到 $V_\pi L_\pi = 2\ \mathrm{V \cdot mm}$. 与文献[28]相比，文献[29]的报道具有更高的调制带宽和更小的波长敏感性.

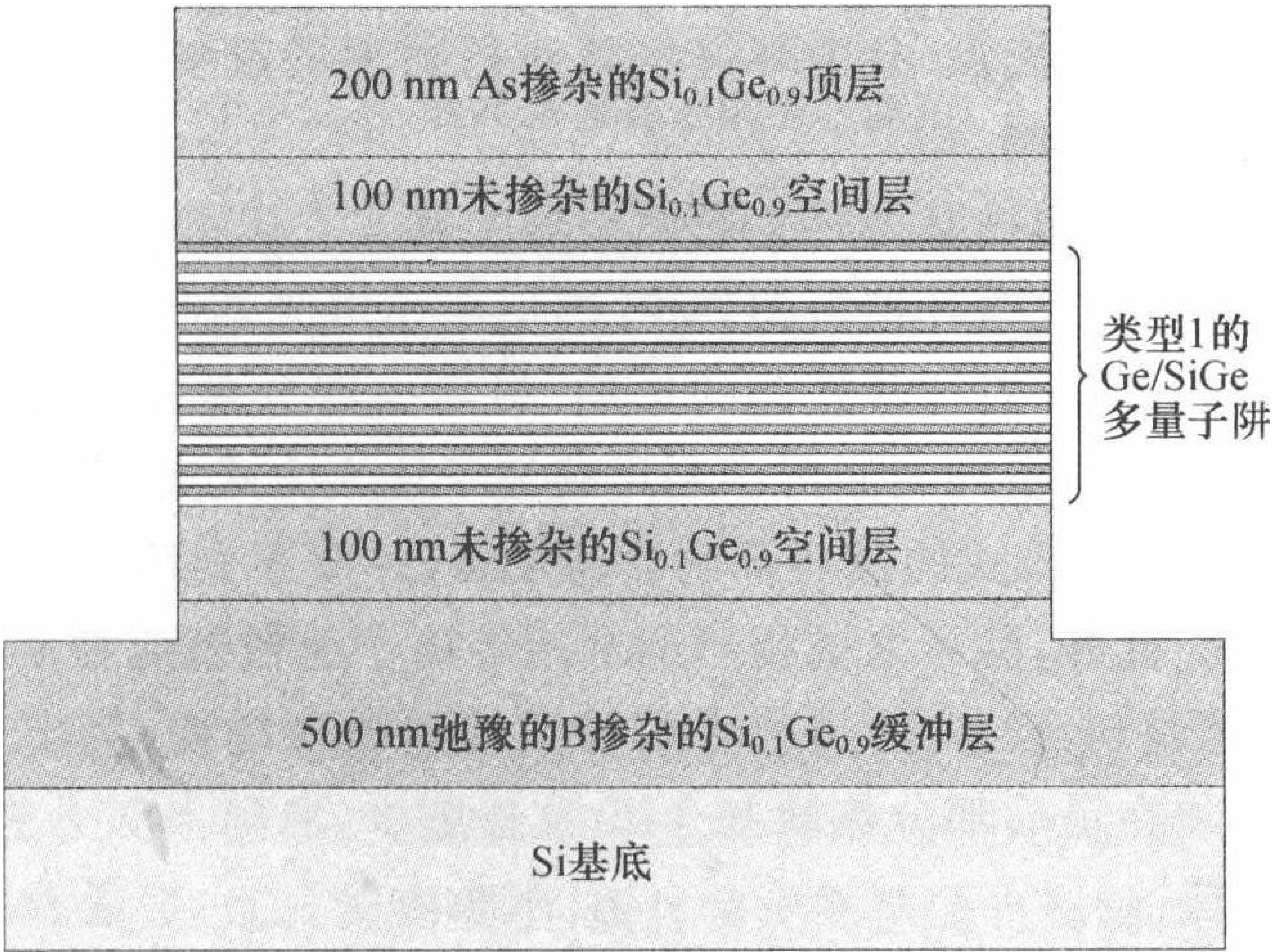


图 6.11　硅基量子阱结构高速电光调制器示意图

5. 小结

自从 2004 年 Intel 公司将调制速度提高到 1 GHz 数量级后，硅基 M-Z 电光调制器的研究获得了一个快速的进展. 但这些进展关注的主要是调制器的速率和调制效率的提高，随着这两个主要性能指标达到商用要求，小尺寸、低插入损耗、超低功耗、与 CMOS 工艺兼容等需求就变得越来越急切. 然而，调制器的这些性能指标的获取通常是相互矛盾的. 例如，调制速度通常和调制效率负相关，为了获得更高的消光比，MZI 臂长通常达到毫米量级，长的臂长就伴随着插入损耗、成本和功耗的增大. 这意味着，尽管在某些性能指标的研究上取得了重要的进展，但要想让硅基电光调制技术得到广泛应用之前，仍然有很多重要的困难需要克服. 我们相信，不久的将来，硅基 MZ 电光调制器将会一步步走向实用.

6.3.4 硅基微环电光调制器

由于硅本身的电光效应较弱，不能高效地改变折射率，所以需要较大的光学结构尺度以实现电光调制. 2004 年，Intel 公司制作了基于 MOS 结构的 M-Z 干涉式调制器，首次实现 1 GHz 的硅基电光调制器，具有重要意义. 但由于载流子浓度改变区域与光模式分布区之间耦合较少，需要较长的调制长度. 后来经过不断的改进，MZI 式的硅基电光调制器已能达到 40 Gbit/s 的调制速度，但器件尺寸仍在几百微米甚至厘米量级，无法满足光电集成对于器件尺寸小型化的需求.

为解决这一问题，研究者们提出了各种新型的光学结构. 其中，微环谐振腔不需要反射端面或光栅等结构来提供反馈，结构紧凑，加之 SOI 高折射率差材料波导的实现，进一步减小微环尺寸，目前报道的微环最小半径已达到 1.5 μm，利于实现高密度的光电子集成. 此外，微环谐振腔尺寸小，其有源器件具有驱动电流小、调制频率高等优点. 且微环谐振峰对折射率的改变非常敏感，波长啁啾小，适合实现高速光调制.

本小节将首先介绍硅基微环电光调制器的基本工作原理，包括电光效应、载流子浓度的改变方式及微环谐振器理论. 然后给出电光调制器的主要性能参数，并分析微环结构的引入对这些性能的影响. 最后一部分是硅基微环电光调制器的最新进展. 由于目前微环调制器多利用载流子注入方式实现，因此将着重介绍 p-i-n 结构的载流子注入式的微环调制器的发展，并简要介绍其他电学结构的微环调制器的情况. 此外，除了之前提到的利用等离子色散效应控制波导介质折射率实现硅基电光调制，近年来，研究者们特别针对微环结构，提出了一种新的电光调制机制，也将在最后部分作简要介绍.

1. 基本原理

(1) 微环谐振器的基本原理.

类似 F-P 腔，微环谐振器是一种光学谐振器件[30]，而它更为突出的特点

则是具有良好的波长选择性、腔内增强特性及高品质因数. 作为一种光学基本单元,微环谐振器具有非常广泛的应用领域,如滤波器、光缓存、传感器、激光器、调制器等. 近年来,随着微纳光电子器件制作工艺的发展、SOI 等高折射率差结构波导的引入,使得微环谐振器的尺寸大大减小,现在已经实验制成了半径为 1.5 μm 的微环. 器件小型化正是目前光电子器件的重要发展方向,有利于实现微纳光电子集成. 对于调制器等有源器件来说,由于微环谐振腔尺寸小,因而具有驱动电流小、调制频率高等优点.

微环谐振器通常包括直波导与闭合环形波导两部分,如图 6.12 所示,光由输入波导的输入端入射,传播至直波导与环形波导相邻处,部分以倏逝波的方式耦合进入环形波导. 类似 F-P 腔,如果光在环形波导中传播一圈改变的相位正好等于 2π 的整数倍,满足相干条件,光波会与新耦合进入微环的光波相互干涉产生谐振增强效应,就会产生谐振. 公式如下:

$$2\pi R n_c = m\lambda, \tag{6.3.12}$$

其中 R 为微环半径,n_c 为微环中光模式的有效折射率,λ 为谐振波长;m 取整数,代表谐振级次. 满足谐振条件的光留在环形波导中,而不满足谐振条件的光从输出波导耦合输出.

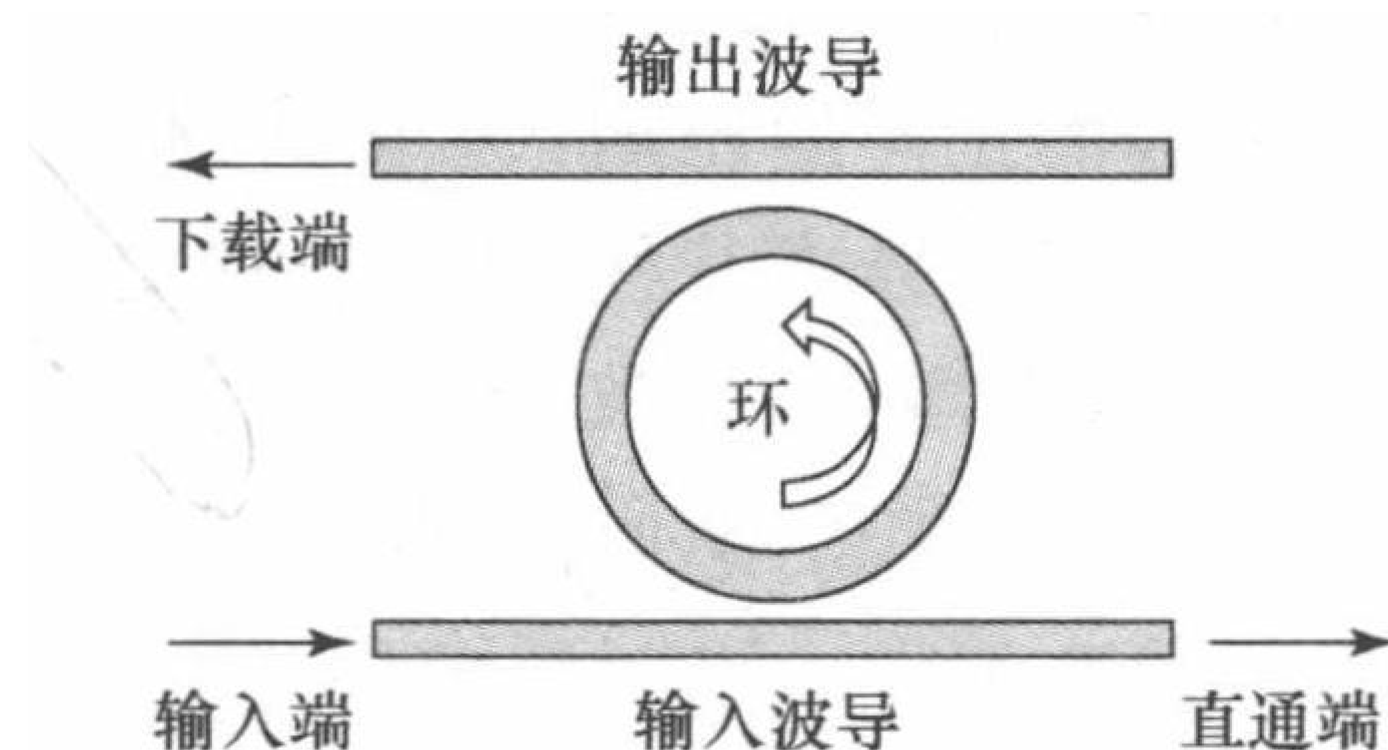


图 6.12　微环谐振器的俯视平面图

微环谐振器的特性主要是微环波导与信道波导之间的耦合系数,可以通过调整微环波导与信道波导之间的间隙来控制.

(2) 微环调制器的工作原理.

由微环谐振条件(6.3.12)经过变化可得谐振波长

$$\lambda = \frac{2\pi R n_c}{m}. \tag{6.3.13}$$

可见,谐振波长与波导的有效折射率成正比关系,利用电光效应改变微环折射率,即可使谐振波长发生漂移,实现微环电光调制器. 很小的折射率改变就可以导致显著的谐振峰偏移,且波长啁啾小,适合实现高速的光调制.

图 6.13 所示为一个基于 p-i-n 结的硅基微环电光调制器[28]. 微环部分由 p-i-n 脊形结构波导构成,中间部分是本征硅作为波导. 两边分别为 p 型和 n 型重掺杂

区域. 该器件通过载流子注入机制实现电压对载流子浓度的调制.

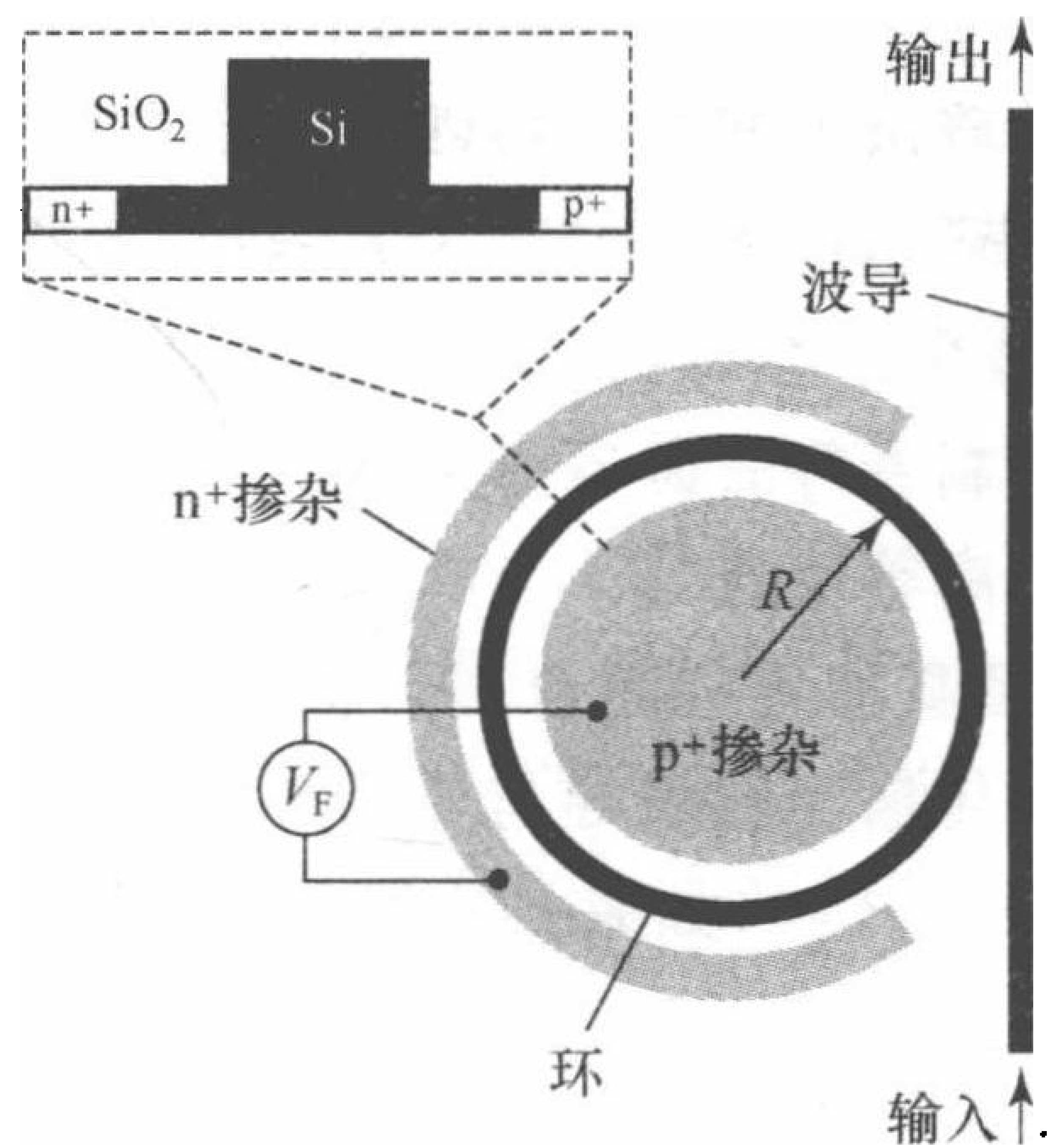


图 6.13 硅基微环电光调制器结构示意图

图 6.14 展示了当 p-i-n 结加不同大小的偏压时,波长在 1574 nm 波段的微环谐振器的透射率的改变. 当所加正向电压为 0.58 V 时,远小于结的内建电场,此时通过结区的电流非常小,载流子浓度基本不变,微环谐振腔保持原有的透射光谱图样. 随着在 p-i-n 结所加的正向偏压增大,波导(本征区)中的载流子浓度增加,有效折射率减小,微环的谐振频率蓝移. 所加正向电压越大,谐振波长的凹陷更浅一些,这是因为载流子浓度增大,导致光损耗增大. 当给器件加不同电压时,某一固定波长处的透射率发生改变,从而实现电信号到光信号的转换.

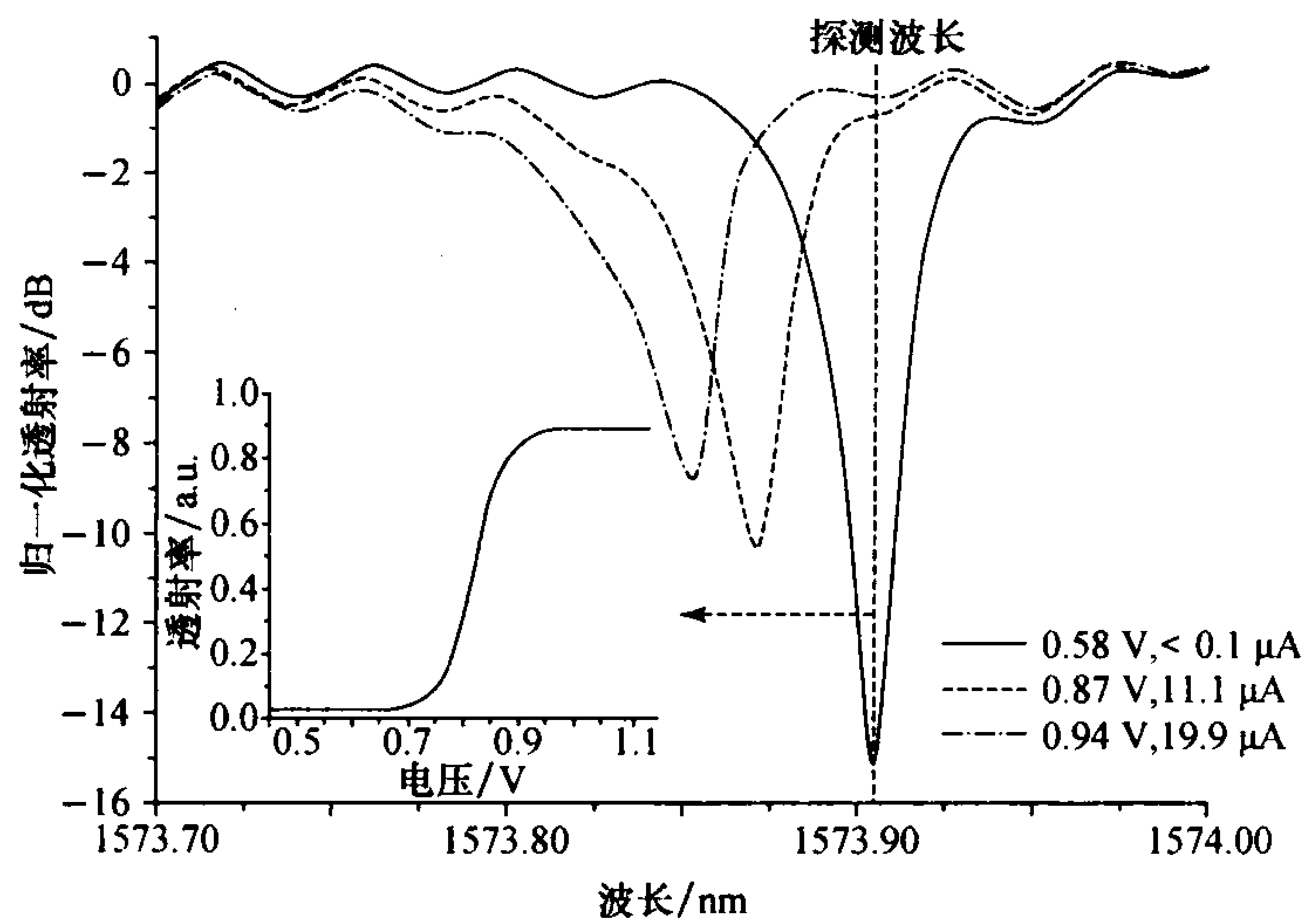


图 6.14 实验测得的透射率频谱

2. 微环电光调制器的主要性能及微环结构带来的影响

电光调制器有多个性能参数，包括调制速度和调制深度，器件损耗和器件功率损耗等. 由于目前多数微环调制器采用的是载流子注入方式(p-i-n 结)，本节将着重分析微环结构对 p-i-n 式的电光调制器各性能参数的作用与影响.

(1) 调制速度.

对于基于等离子色散效应的硅基电光调制器来说，调制速度主要由载流子浓度的改变速度决定. 与其他结构相比，载流子注入式的 p-i-n 结构一般被认为最不易实现高速调制，但微环谐振腔的引入却能够对这一性能有一定的改善. 在 p-i-n 结构中，载流子的抽运是通过反向偏压条件下的漂移运动实现的，速度可达几十皮秒量级，而真正限制调制速度的是载流子的注入过程，它通过载流子的扩散实现，通常需要 10 ns 数量级的时间. 然而，由于微环谐振腔的谐振波长对折射率的变化非常敏感，实验数据表明[9]，当所加电压较高时，在 p-i-n 结中载流子浓度的变化达到稳态之前，光透射率就非常接近最大值 1 了. 这意味着光信号的上升时间可以远小于电信号的上升时间(约 10 ns)，这对实现高速调制具有重要意义.

(2) 调制深度.

对于光互连等应用，通常要求调制深度在 7 dB 以上，而对整体系统链路预算来说，4～5 dB 的调制深度就足够了. p-i-n 结载流子浓度变化区与光场分布区耦合较大，加之微环谐振频率对波导折射率的变化非常敏感，因此利于实现大的调制深度.

(3) 器件损耗.

器件损耗可包括由于调制器的接入所带来的插入损耗以及光信号在调制器中传输的损耗. 插入损耗通常由反射、吸收以及模式耦合损耗等组成. 而光信号在调制器中传输时的损耗则主要来源于波导对光的吸收等. 事实上，硅波导在通信波段对光的吸收比较小，SOI 波导单模条件下，损耗最低可小于 0.01 dB/cm. 然而在硅基电光调制器中，由于载流子的存在，会产生自由载流子吸收损耗，比材料本身所产生的损耗大得多，一般大于 1 dB. 因此在器件设计时，要格外注意重掺杂区与光模式存在区域的距离. 例如在 p-i-n 结构，掺杂区距离微环硅波导接近 1 μm，就是为了尽量减小光谐振模式的空间分布与掺杂区之间的耦合，以减小吸收损耗.

(4) 器件能耗.

相较于基于 MZI 等结构的硅基电光调制器，微环电光调制器具有最小的尺寸，有望实现低能耗. 然而，由于微环谐振器的谐振峰宽度非常窄(单环结构约 0.1 nm，高阶环结构为 0.5～2 nm)，加之硅材料的热光效应系数较大，微环电光调制器对温度非常敏感. 因此，为了保证工作温度的稳定，需要额外增加热电温

度控制装置[29]，增大了整个器件的能耗.有的文章提出，多环级联的方式可以减弱温度漂移问题[31]，但只适用于较窄的光频范围.

3. 研究现状

硅基微环电光调制器的研究历史并不算长，2005 年由美国康奈尔大学 Lipson 教授的研究组首次提出.随后的几年，这一领域一直是国内外的研究热点，除了康奈尔大学、美国南加州大学，北京大学、浙江大学、上海交通大学、中国科学院半导体所等都在积极开展相关研究.

(1) 基于载流子注入方式(p-i-n 结)的硅基微环电光调制器.

以 Lipson 教授组的研究工作为代表，目前大多数硅基微环电光调制器仍采用的是 p-i-n 结构.如前所述，这种电学结构的最大缺陷在于难以实现较高的调制速度，因而相关的研究工作主要集中在提高调制速度方面.此外，继续减小器件尺寸及器件功耗也是重要的研究方向.

图 6.15 所示是 SEM 下所观察到的实验结构.微环直径为 12 μm，比当时的其他结构的硅基电光调制器在尺度上减小三个数量级.调制速率为 1.5 Gbit/s.虽然器件尺寸有了显著减小，但调制速率与同时期的其他硅基电光调制器相比还有一定差距.Intel 公司也已制成调制速度达到 10 Gbit/s 的 MOS 结构电光调制器(但仅有源区长度就长达 4.5 mm).

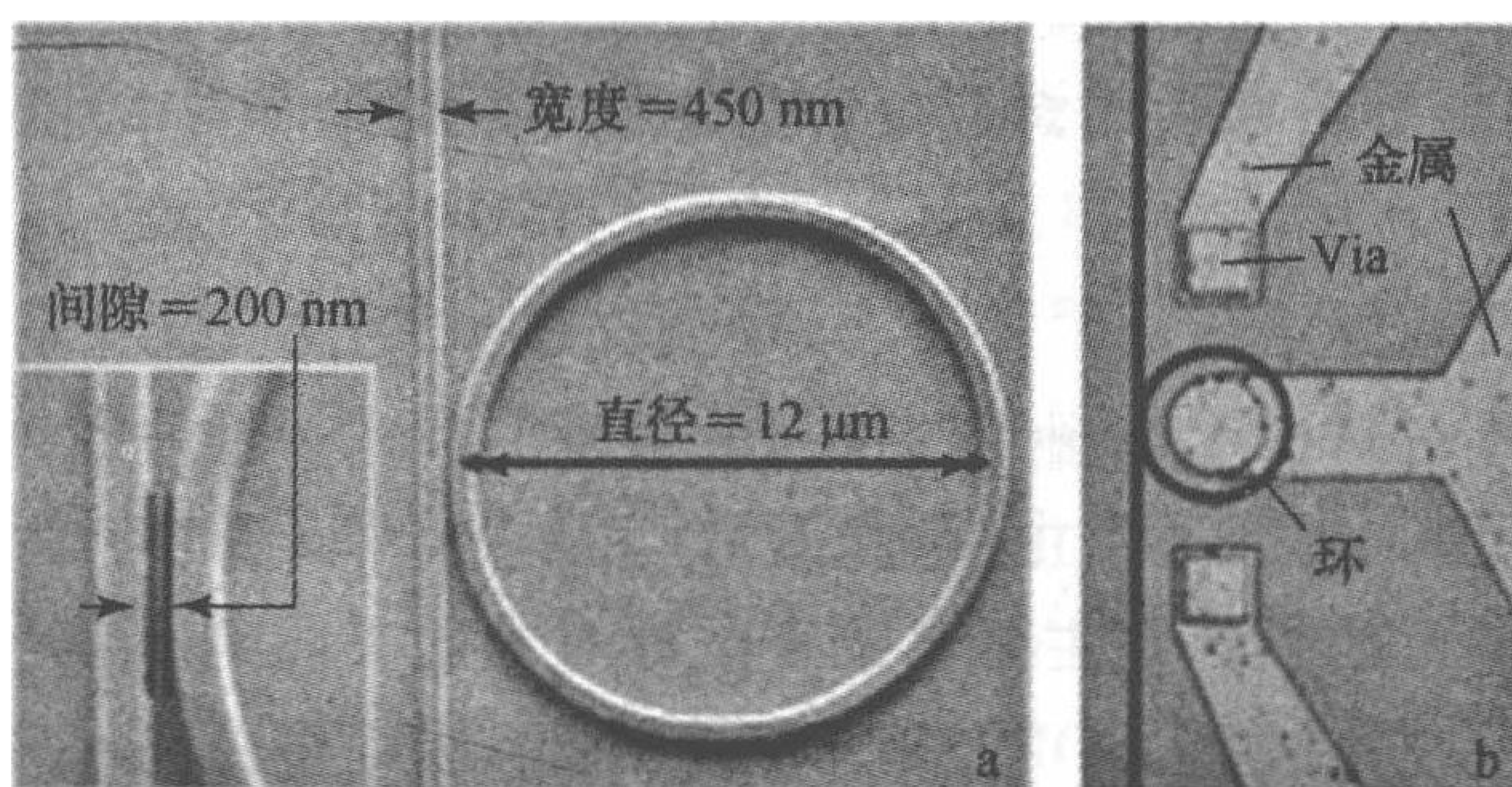


图 6.15 SEM 下的实验结构

2007 年，研究者对器件进行了改进和优化[32].与之前的结构相比，微环直径减小到 5 μm，在直波导另一侧加了一块 n 型重掺杂区域，在耦合区域也能形成近似 p-i-n 结的结构.并将 p-i-n 结构的重掺杂区域与本征区波导之间的距离减小到近 300 nm.此外，更重要的一个突破是使用了预强调驱动信号，进一步减小了载流子浓度上升与下降的响应时间，将调制速度提升至 12.5 Gbit/s，如图 6.16 所示.在硅基材料中使用预增强驱动信号的方法是由 Png 等[33]于 2004 年首次提出的，这种方法基于光信号响应速度小于电信号响应速度的事实，通过巧妙的控制驱动信号的形状来减小器件达到稳态的时间，提高调制速度.

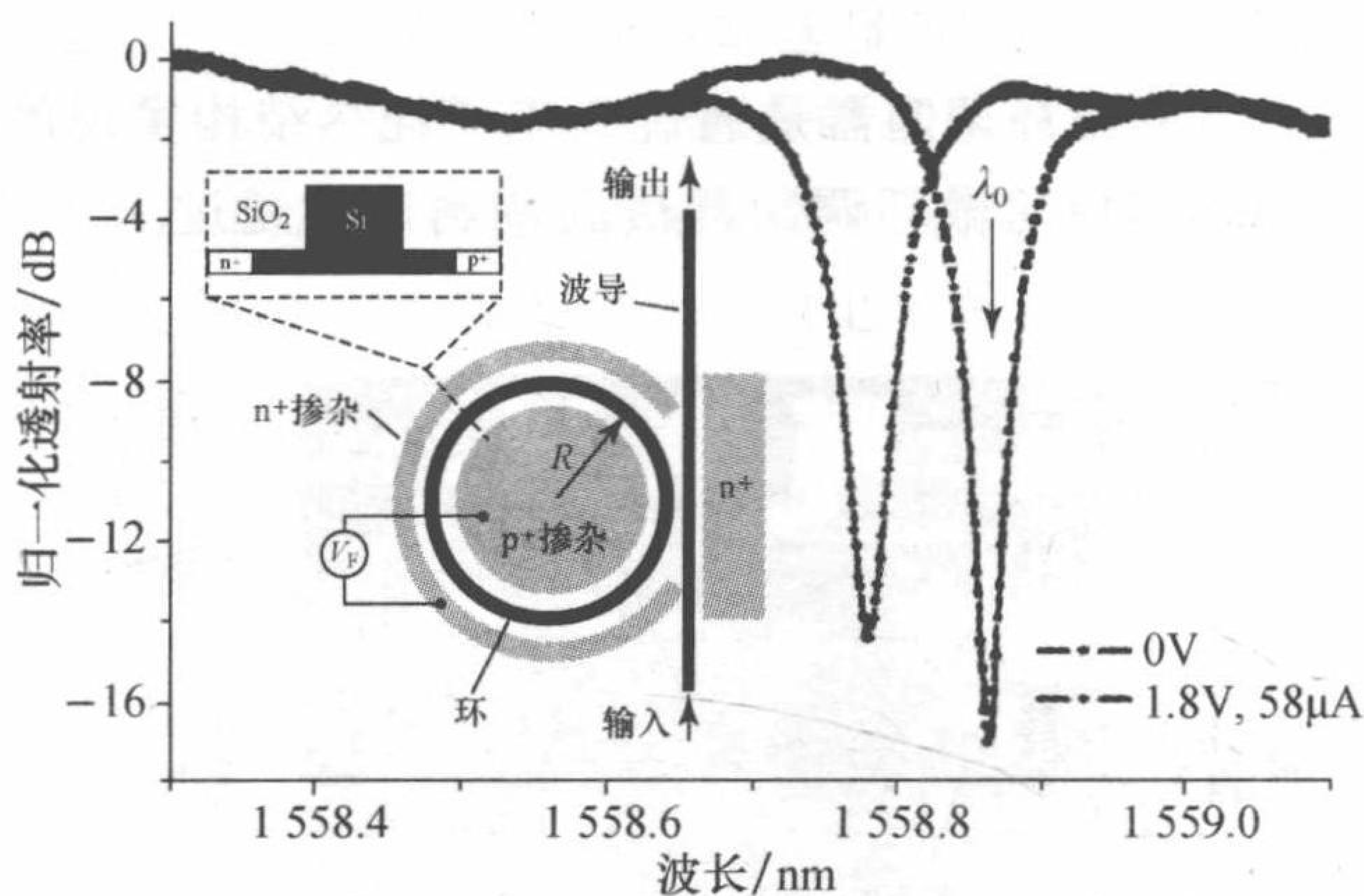


图 6.16　12.5 Gbit/s 载流子注入式硅基微环电光调制器[9]

调制器的速度是非常重要的性能之一，但在实际应用中，特别是光互连等光电子集成领域，器件功耗也是很重要的性能指标. 2010 年，研究者在保证一定的调制速度(1 Gbit/s)的前提下针对驱动电压这一性能对器件进行优化，实现了峰峰值 150 mV 的超低驱动电压[34]，这是目前的硅基调制器在 Gbit/s 量级的调制速率能实现的最小驱动电压. 如图 6.17 所示，器件结构仍是基于 p-i-n 的单个微环与单个直波导构成的微环谐振器，通过载流子注入的方式改变载流子浓度. 器件尺寸有进一步的减小，模式体积($2\ \mu m^3 \sim 0.52\lambda^3$)和占用面积($20\ \mu m^2$)是目前硅基微环调制器的最小尺寸. 实现超低驱动电压的关键技术在于合理选择电荷注入时机. 通过研究电压改变值与电荷注入量的关系，发现在偏压为 1V 左右有一个最佳的电荷注入区域，此时 100 mV 的电压改变量就可以引起 2×10^{17} 数量级的载流子浓度改变，相应的波导折射率变化为 7.7×10^{-4}. 纳米光电子调制器驱动电压的减小进一步促进了未来低功耗微纳光电子集成的实现.

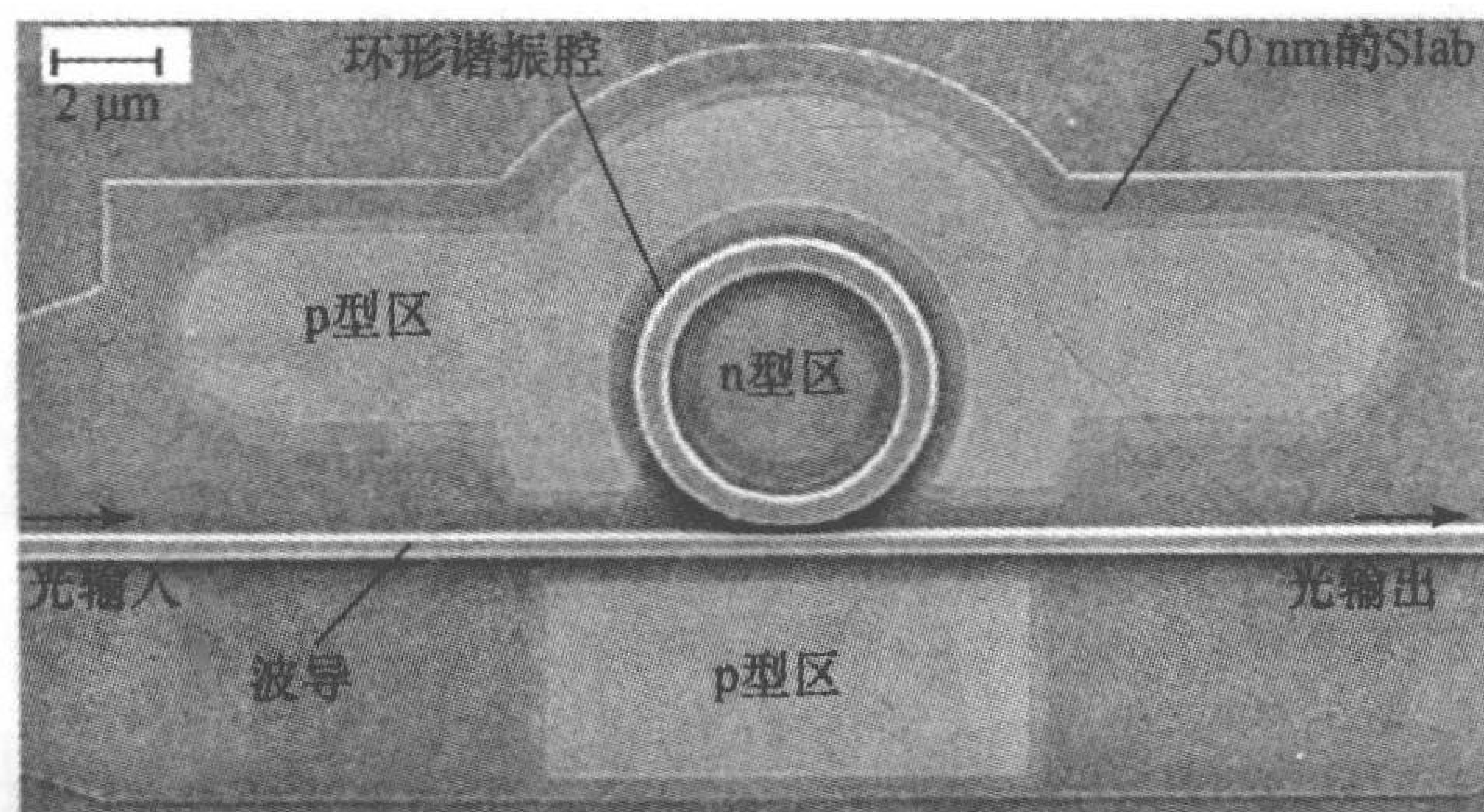


图 6.17　超低驱动电压、超小模式体积的硅基微环电光调制器[34]

(2) 基于载流子耗尽机制的硅基微环电光调制器.

除了 p-i-n 结,有的微环调制器是通过载流子耗尽结构实现的.目前报道的基于载流子耗尽原理实现的微环调制器最高调制速度超过了 35 GHz[35],如图 6.18 所示.

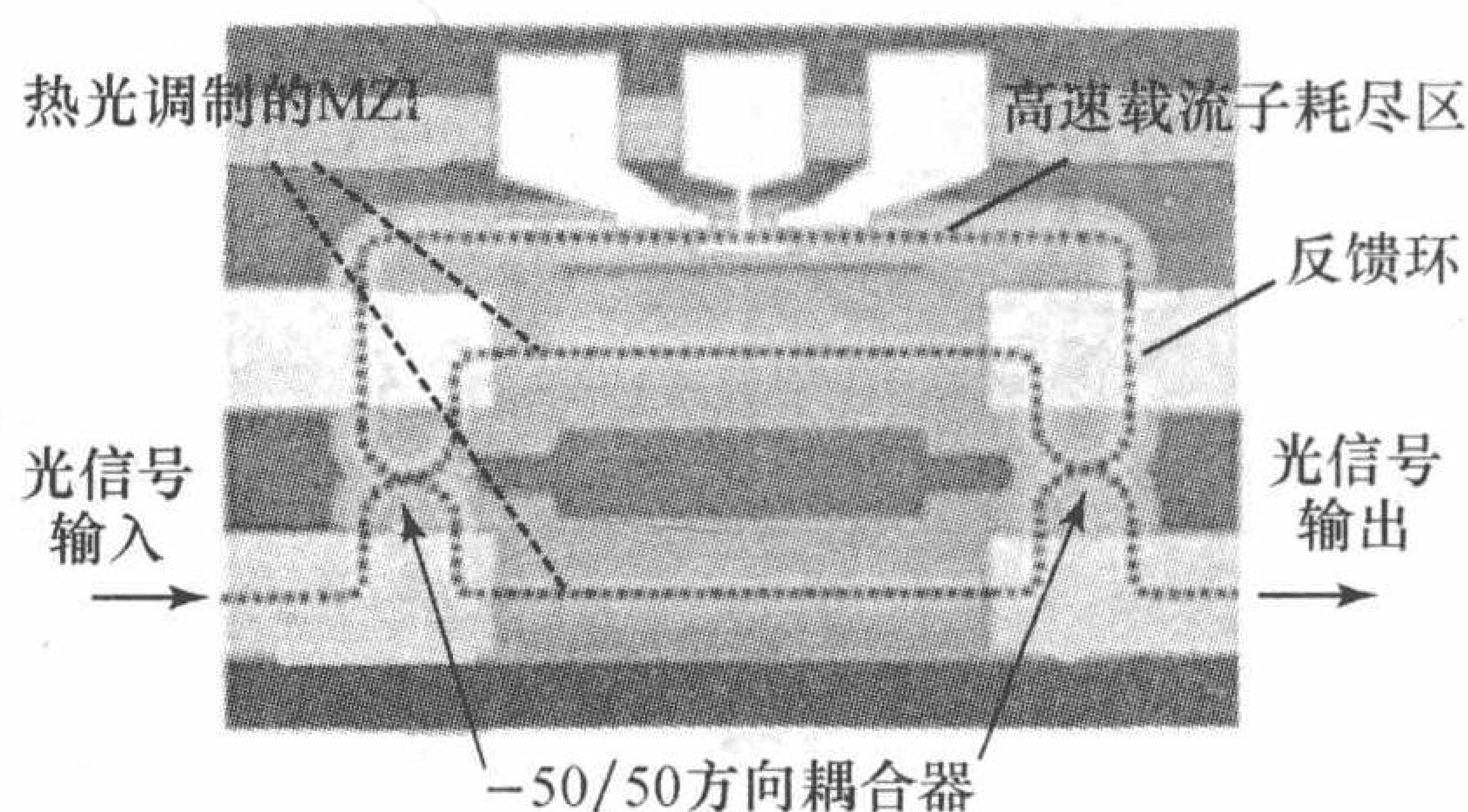


图 6.18　具有热光调制对称性 MZI 结构的高速载流子耗尽式微环电光调制器

(3) 其他调制机制.

利用等离子色散效应通过电注入控制波导介质折射率是目前实现硅基电光调制器的一种最普遍的方式.这种方法适用于多种光学结构,包括 MZI、布拉格光栅、微环等,分别结合不同的电学结构及电驱动信号方式,满足不同的性能需求.然而,对于本文主要讨论的微环结构,近年来有学者提出了独特的调制机制[36].

如图 6.19 所示,在微环谐振腔中嵌入了一个 MZI 结构,MZI 与微环波导通过耦合器相连.通过推挽式电压信号控制光信号在 MZI 的传播特性,从而改变微环谐振腔与直波导间的耦合系数,进而实现对光信号的调制,解决了传统调制方式的啁啾问题.这种结构早有报道,但将其应用在电光调制器方面是由上海交通大学的叶通在 2009 年才首次提出的[36].随后,作者又对这个结构进一步改进[37],减小其功耗,如图 6.20 所示.

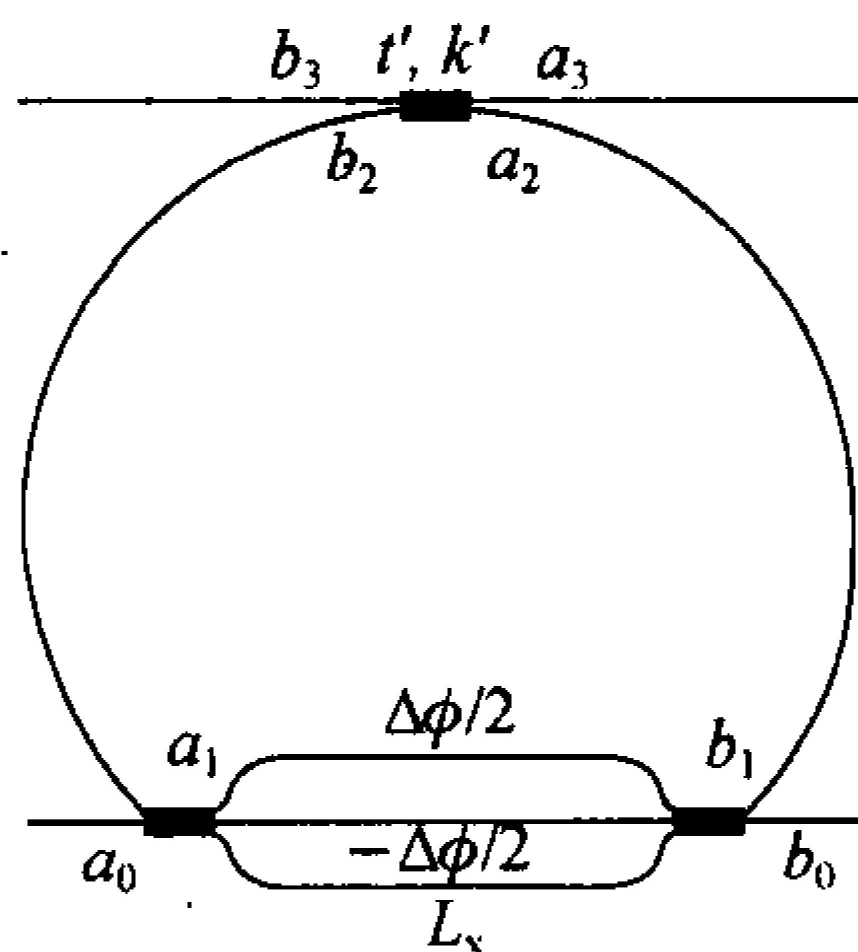


图 6.19　推挽式微环调制器

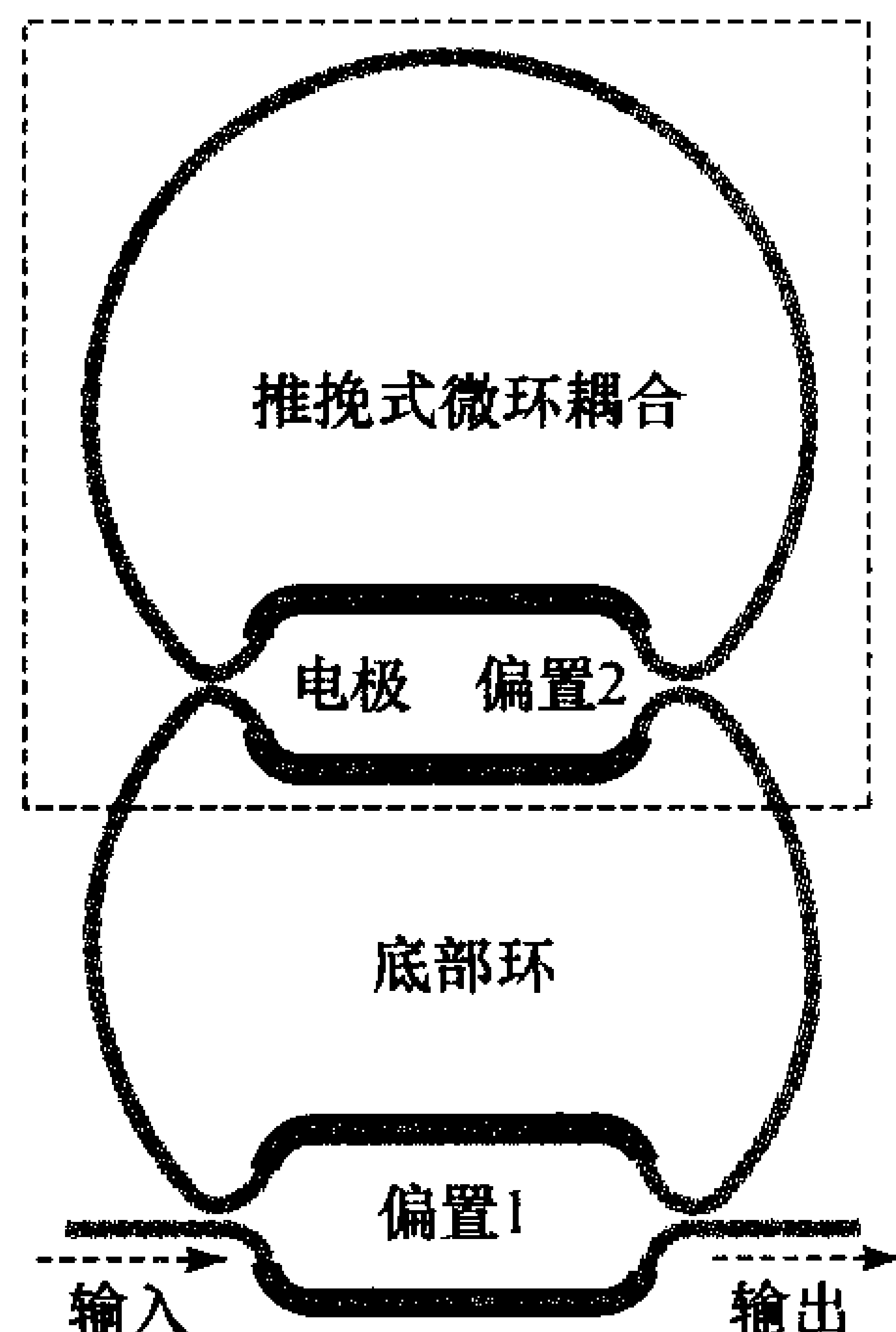


图 6.20　双环结构

4. 问题与展望

虽然微环结构的引入使硅基电光调制器的某些性能带来显著改善,同时也有许多问题. 与 MZI 调制器相比,谐振条件的限制使得谐振腔结构的调制带宽小得多. 例如,在 MZI 中能实现 20 nm 以上的结构在微环谐振腔的情形下只有大约 100 pm. 窄带器件的应用会受到一定的限制. 此外,微环结构对工艺制作精度要求较高. 例如,微环波导的宽度平均每改变 1 nm,就会使谐振频率改变近 0.25 nm. 微环谐振频率具有很高的温度敏感性,而硅材料本身的热光效应比较明显,因此硅基微环电光调制器需要额外的热电冷却/控制器以保证器件工作温度的稳定,这会增加整个器件的能量消耗,违背了减小器件耗能的初衷.

通过回顾硅基微环电光调制器的发展历程可以发现,无论是选用不同的电学结构还是采用不同的电信号驱动方式,提高器件性能的工作主要集中在电学性能方面的研究与改进,这正是限制光电子器件各方面性能进一步提高的主要问题. 新型光学结构(如多环级联)与新的调制机制的应用则为微环调制器的发展注入新的血液,具有广阔发展前景.

§ 6.4 硅基热光调制

硅基光波导具有良好的光学传输性能,同时又与传统的硅加工工艺兼容,因此近来在光电子器件及集成方面得到广泛的应用. 由光波导构成的强度调制器是集成光电子学中的重要器件之一,广泛应用在高速数模转换、光学逻辑回路、光传感技术等方面. 但由于硅是中心对称的晶体,直接电光效应很弱,只能通过等离子色散效应[4]、热光效应和声光效应来进行折射率调制.

基于等离子色散效应的电光调制[4]速度快,其缺点是注入的高浓度载流子会产生载流子吸收,在调制过程中带来能量的损耗,而且由大电流注入引起的发热会降低电光调制的效率. 硅基热光调制速度相对较慢,但由于硅的热光系数很高,它在低成本和低频调制领域具有很大的吸引力. 本小节主要对国内外硅基热光调制的基本原理、器件结构、性能参数进行阐述,有关声光调制部分将在下节作详细介绍.

6.4.1 硅基热光调制原理和结构

众所周知,温度升高会引起材料膨胀和极化. 而硅材料的热膨胀系数相对极化系数来说很小,因此,硅的折射率受材料极化影响较大. 我们把硅材料的折射率随温度变化而变化这种现象称为热光效应[38]. 根据硅材料的特性,其热光效应可以简单的表达为

$$\frac{dn}{dT}=1.86\times10^{-4}/K. \tag{6.4.1}$$

由上式可知，如果温度上升 6℃，硅折射率变化为 1.1×10^{-3}. 这个变化已经可以与等离子体色散引起的电光效应相媲美了. 因此，只要设计制作一种硅基光子器件，使其能利用任何能量形式使硅材料的温度发生变化，并对通过的光信号进行调制，这也就是最基本的硅基热光调制器.

硅基热光调制器的结构很多，但主要有 MZI 干涉仪型、光子晶体结构型、微环型等等，下面我们主要对这三种结构进行介绍.

1. MZI 干涉仪型

干涉型主要以 Y 分支 MZI 热光调制器为主，其结构如图 6.21 所示.

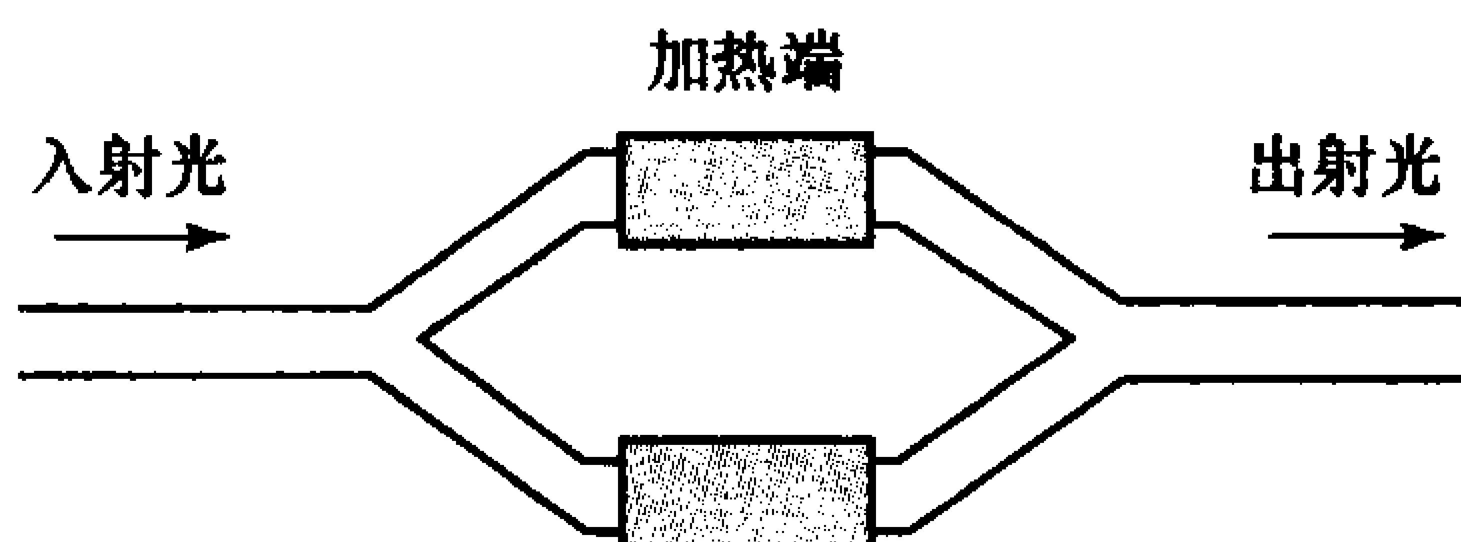


图 6.21 MZI 干涉型热光调制原理图

这种结构基于 SOI 的硅基波导，在硅波导上加一层二氧化硅层，再在二氧化硅层上加一个金属电极加热，热通过二氧化硅层再传到硅层，改变硅的折射率，从而实现热光调制. 但是由于二氧化硅的导热性较差，所以热从二氧化硅层传到硅层很慢，这样就使得调制速度很低. 改进的办法是将加热电极层直接加到硅上，这样的边加热电极结构使得调制器的速度大大提高. 由热光效应引起的光的相位变化与调制器的臂长以及硅温度的变化可以由下面的公式表示：

$$\Delta\varphi=\frac{2\pi}{\lambda}\left(\frac{dn_{eff}}{dT}\right)\Delta TL, \tag{6.4.2}$$

其中 λ 是工作波长，n_{eff}是波导的有效折射率，ΔT 是硅材料的温度变化，L 是调制臂的臂长，$\frac{dn_{eff}}{dT}$大约等于材料的热光系数. 对于确定的调制器，L 一定. 只要知道温度的变化，就可以得出相位的变化. 在输出端两个臂产生干涉从而实现光调制.

这种热光调制结构紧凑，制作容差大，偏振不敏感等优点，成为热光调制的热点.

2. 光子晶体结构型

光子晶体是具有 PBG 特性的人造周期性电介质结构，PBG 是指某一频率范围的波不能在此周期性结构中传播，即这种结构本身存在“禁带”.

光子晶体型热光调制如图 6.22 所示，其中中间是光波导，浅白色区域是加热区域，当光通过加热区域时，由于热光效应导致该区域折射率发生改变，从而引起禁带的截止波长的改变，实现调制功能[39]. 其优点为截止波长变化范围大、调制的频率范围广、消光比高、插入损耗小，缺点为制作工艺复杂.

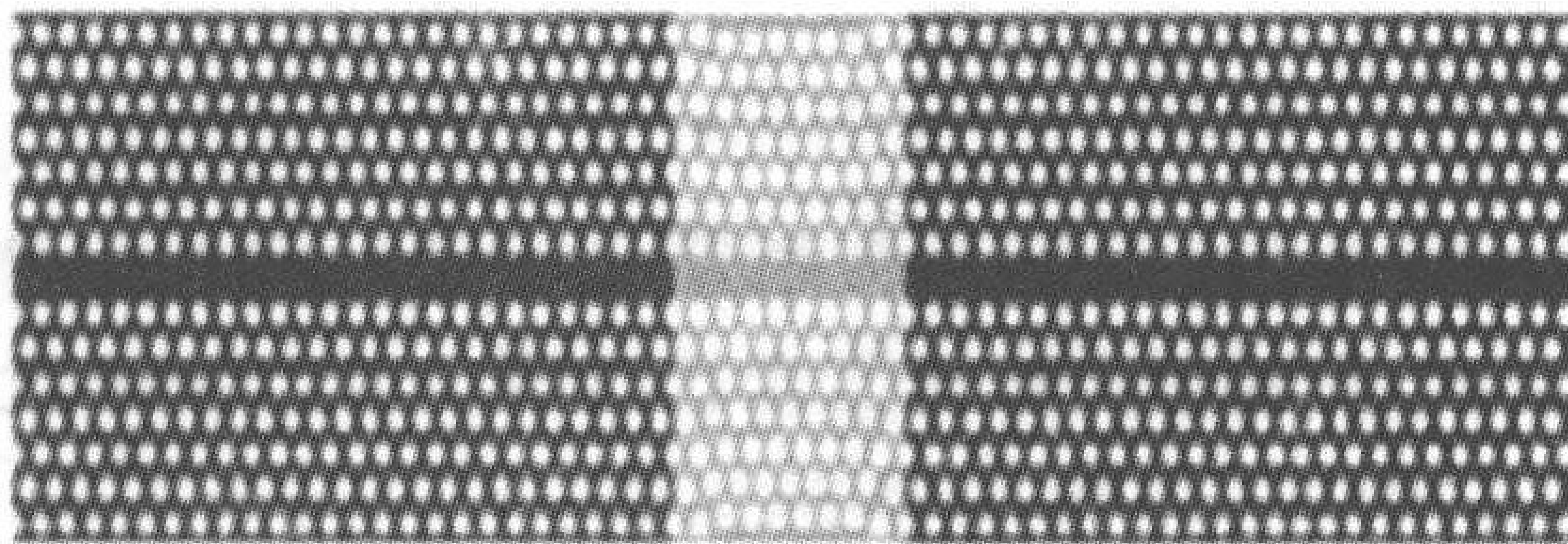


图 6.22 光子晶体热光调制原理图

3. 微环型热光调制结构

微环型热光调制器[40]原理同样是由于温度引起折射率的改变,从而改变微环谐振腔的谐振波长,达到调制的作用.微环谐振方程

$$\frac{2\pi}{\lambda}RN_{\mathrm{eff,ring}}=m,\tag{6.4.3}$$

其中 λ 是光波长,R 是环半径,$N_{\mathrm{eff,ring}}$ 通过环的光模式有效折射率,m 是谐振级次,温度改变引起折射率的改变,从而引起有效折射率的改变,使得谐振波长漂移,漂移量[41]为

$$\delta\lambda_{\mathrm{C}}=\frac{\lambda\Delta N_{\mathrm{eff}}}{N_{\mathrm{eff,ting}}}.\tag{6.4.4}$$

其原理与微环电光调制类似,差别在于引起折射率改变的外界能量.微环型热光调制器由于其低驱动电压,高调制效率,高调制速度等优点而被广泛研究,缺点是制作工艺要求高,对环境影响敏感.

6.4.2 主要性能指标

1. 调制速率

热光调制器的调制速率可定义为温度变化引起光强度变化的反应时间,也就是调制器的响应时间.由于材料温度的变化速度有限,因此热光调制带宽有限,微环型热光调制由于驱动电压低,其调制速率最高.

2. 功耗

功耗是热光调制中一个不可忽视的参数,热光调制的功耗低,效率高,损耗就少,对环境温度的影响小;效率不高的热光调制器所浪费的热功率将会对热光调制器产生影响,降低调制度.

3. 调制度

调制度是调制器中基本的性能参数,其调制度定义为

$$\eta=\frac{P_{\mathrm{m}}-P}{P_{\mathrm{m}}},\tag{6.4.5}$$

其中 P_{m} 是指没有加调制时候的光功率,P 是调制时的最小光功率;η 越接近 1,调制效果越好.

6.4.3 热光调制研究进展

由于热光调制速度的限制，热光调制器只能在低频中应用，如利用 Y 分支来分束/合束 MZI 热光调制器[42]，其消光比达到－16.5 dB，上升时间 10 μs，下降时间 20 μs，实现 π 相移功耗 0.39 W.

光子晶体结构的热光调制由于其大的消光比以及宽的调制范围，一般用此来设计通信波长的截止频率，M. T. Tinker 采用图 6.22 所示的结构在通信波长 1550 nm 处实现截止波长的移动达到 60 nm，消光比达到－50 dB 的光子晶体热光调制，如图 6.23 所示[39].

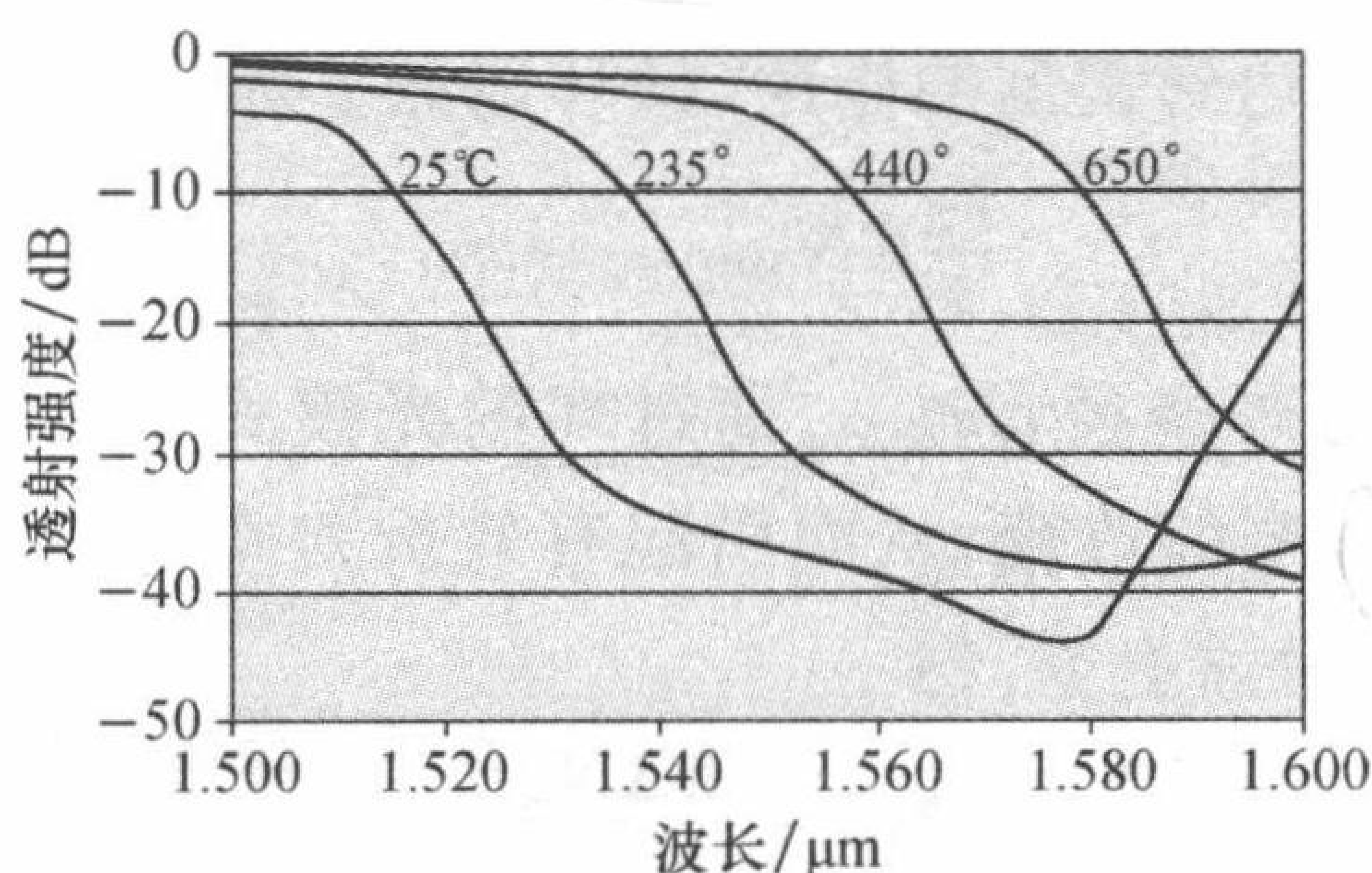


图 6.23 截止波长的漂移与温度的关系

微环由于其高 Q 值及高敏感特性也被应用到热光调制中，微环热光调制可用于微环滤波器，微环开关阵列，也可制作微环调制器. D. H. Geuzebroek 采用微环热光调制功能实现了功耗 10 mW，开关速率 2 kHz 的微环光开关[40]. 2009 年美国 Sandia 国家实验室 M. R. Watts 提出了采用掺杂离子的方法将热极做在微环中的方法[41]，其结构如图 6.24 所示，谐振腔加宽的区域是 n 掺杂形成热源接触点，通过高掺杂的类似于细链的绝热却导电的导电线连接到一个钨丝盘底，再接到外加电极的铝线上，达到电加热的方法，从而改变微环折射率实现热光调制的效果；调谐低能耗能达到 4.4 μW/GHz，调谐的热速度能达到 1 μs[43].

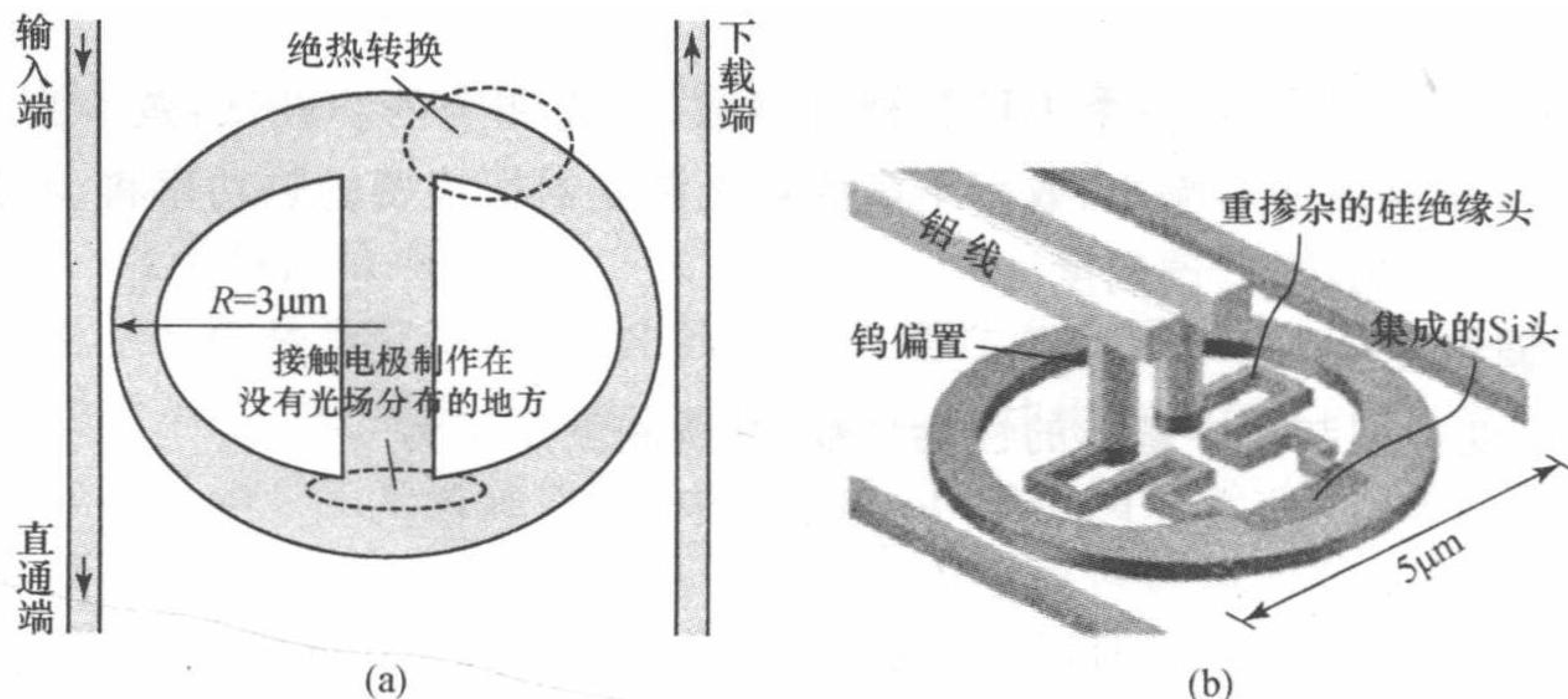


图 6.24 高速热光微环调制器结构原理图

由于硅的热光效应比电光效应引起的折射率改变量大得多，因此可用热光效应来做微环电光调制波长的可调谐. 2010 年，P. Dong 采用图 6.25 结构的方法，在微环电光高速调制的情况下，采用热光效应实现调制波长的移动，实现了波长可调谐长度为 8 nm 的 12.5 Gbit/s 高速可调谐微环电光调制器，驱动电压 3V 时就能达到 −6 dBd 的消光比，调谐效率达到 2.4 mW/nm[44]. 2007 年 F. Gan 利用热光效应实现了最大波长可调谐范围的微环调制器，FSR 为 16 nm，而最大可调谐范围可达到 20 nm，调谐效率为 28 μW/GHz[45].

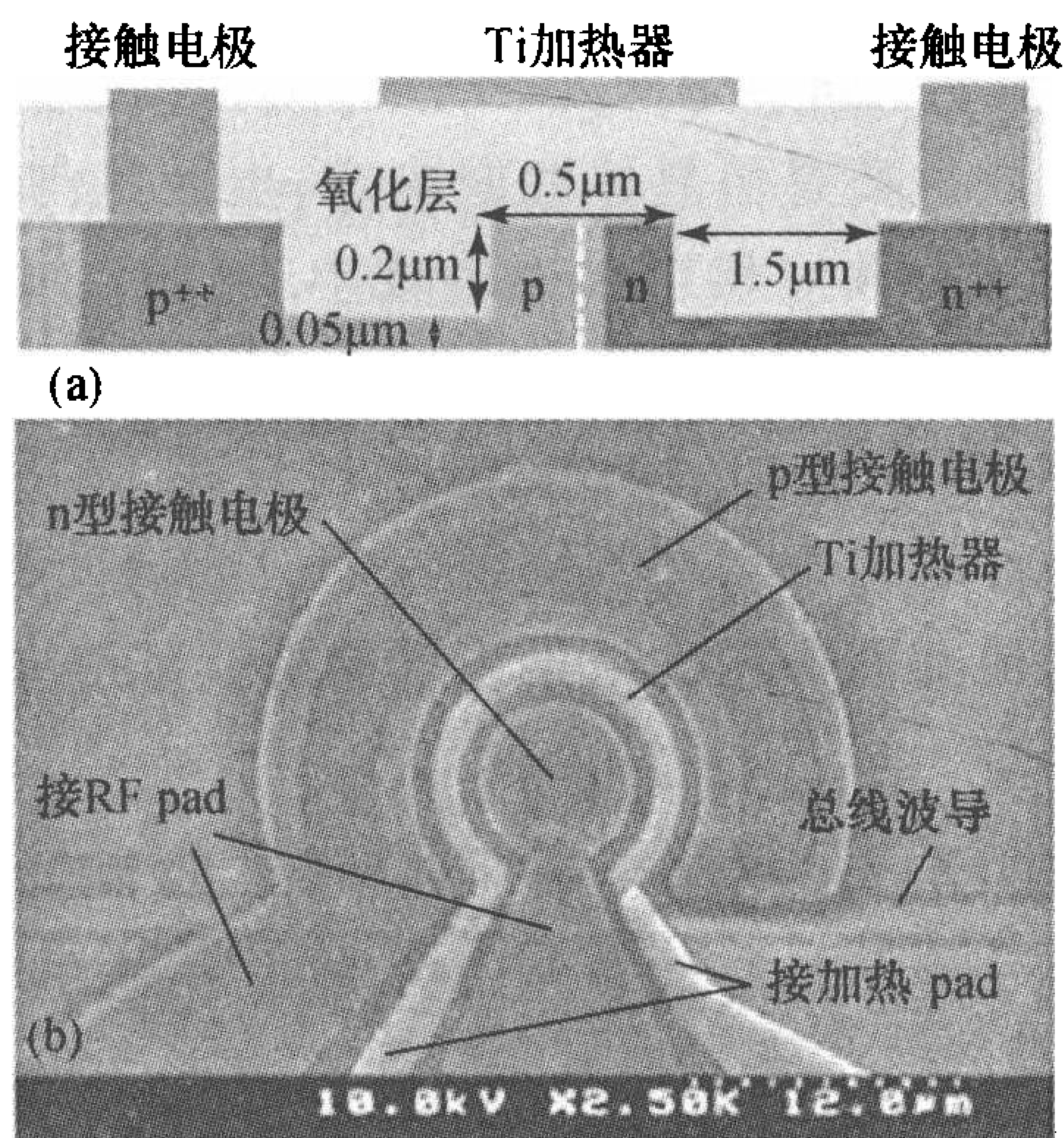


图 6.25 可调谐微环谐振器原理图

6.4.4 小结

通过热光效应介绍了几种结构的热光调制的结构、工作原理及性能参数，阐述了各结构的优缺点，并分别对各种结构的应用进行了介绍，总结了国内外热光调制应用及发展状况.

§6.5 硅基声光调制

声光效应是弹性波与光波的相互作用所表现出来的一种物理现象. 当超声波通过介质时会造成介质的局部压缩和伸长而产生弹性应变，而且该应变随时间和空间作周期性变化，从而使介质的折射率发生改变. 当光通过这一受到超声波扰动的介质时就会发生衍射现象，这种现象称之为声光(AO)效应.

6.5.1 声光调制原理

在超声波作用下的此类介质可视为一种由声波形成的位相光栅(称为声光

栅),其光栅的栅距(光栅常数)即为声波波长.当一束平行光束通过声光介质时,光波就会被该声光栅所衍射而改变光的传播方向,并使光强在空间作重新分布.由弹性应变引起介质折射率的变化与声功率的关系为可表达为[2]

$$\Delta n = \sqrt{n^6 p^2 10^7 P_a/(2Qv_a^3 A)}, \tag{6.5.1}$$

其中 n 是无应力时介质的折射率,p 是介质的弹光张量,P_a 是超声驱动总功率,Q 是质量密度,v_a 是声速,A 是声波场与介质的交叠区域面积,式中全部采用标准单位.定义一个声光系数 $M_2 = n^6 p^2/(Qv_a^3)$,则表达式简化为如下形式:

$$\Delta n = \sqrt{M_2 10^7 P_a/(2A)}. \tag{6.5.2}$$

在晶体中,声光效应主要是由晶体取向决定,也就是与 p 相关.但是即使选择最优的材料和取向,这种效应相对来说也较小.例如:光波长为 6328Å,石英的 M_2 为 $1.51\times10^{-18}\ \mathrm{s^3/(g\cdot m)}$,而 $LiNbO_3$ 的 M_2 为 $6.9\times10^{-18}\ \mathrm{s^3/(g\cdot m)}$,即使声功率密度为 100 W/cm^2,折射率的改变也只在 10^{-4} 数量级.虽然折射率改变较小,但每一个声波峰值都会在不同的空间改变介质的折射率,因此,如果相位条件满足,不同峰值共同作用的结果积累起来就可形成衍射效应,从而对光场产生较大的影响.

光波衍射效应可通过体声波或声表面波(surface acoustic wave, SAW)相互作用产生,由于硅基光波导只有几毫米厚,因此在光集成应用中如声光调制器或声光开光都是用声表面波产生光衍射.

声光衍射主要分为布拉格衍射和拉曼-奈斯(Raman-Nath)衍射两种类型.前者通常声频较高,声光作用程较长;后者则反之.理论上布拉格衍射的衍射效率可达 100%;拉曼-奈斯衍射中 1 级光的最大衍射效率仅为 34%左右,所以实用的声光器件一般都采用布拉格衍射.下面分别对两种衍射型调制器进行分析.

1. 拉曼-奈斯型调制

拉曼-奈斯型调制器的基本结构如图 6.26 所示.

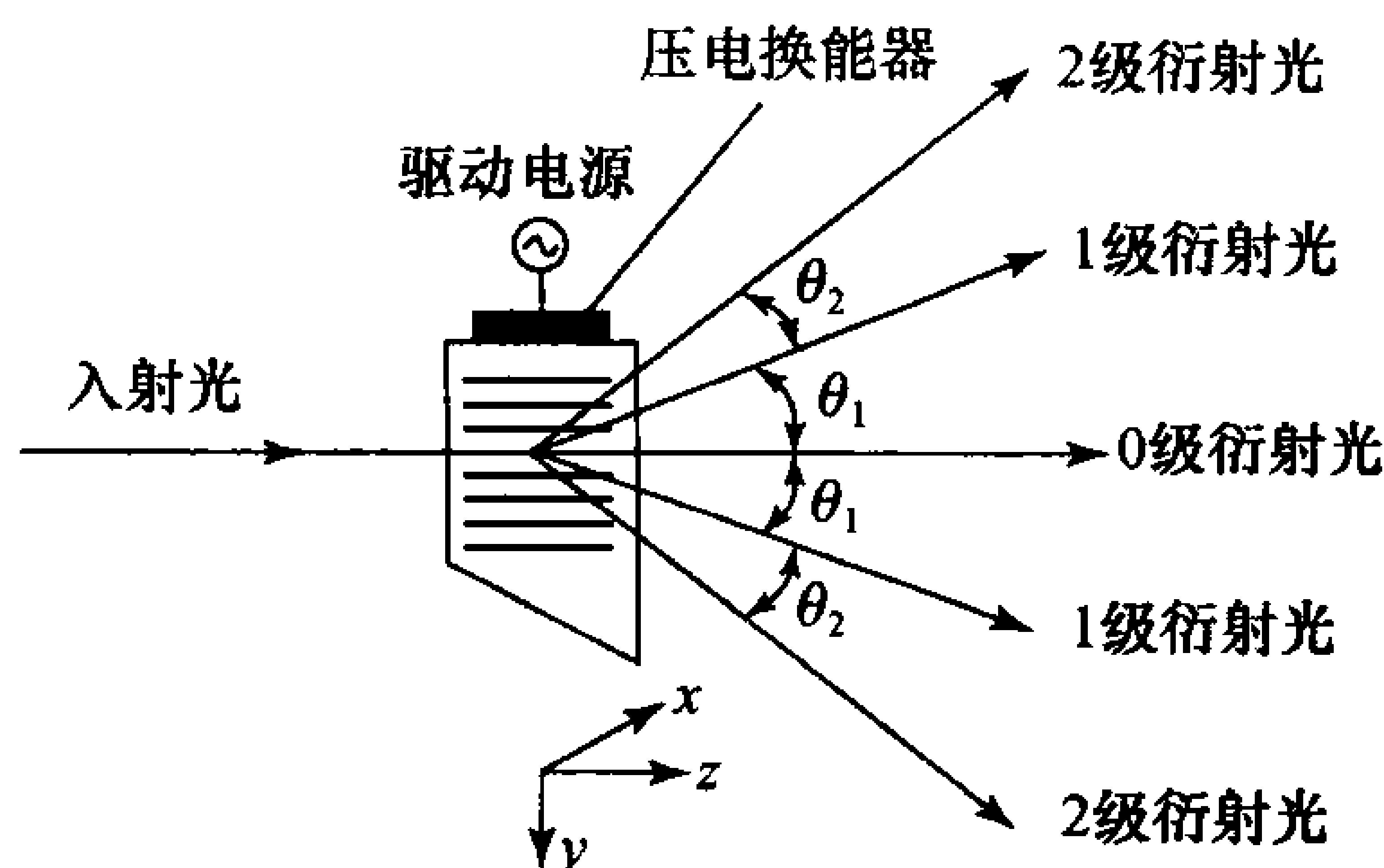


图 6.26 拉曼-奈斯型衍射原理图

光沿 z 方向入射，经过声光通道相位改变为

$$\Delta\varphi = \frac{\Delta n 2\pi l}{\lambda_0}\sin\left(\frac{2\pi y}{\Lambda}\right), \tag{6.5.3}$$

其中 Δn 声光效应引起的折射率改变，l 是相互作用的长度，Λ 是声波波长也就是光栅常数，y 的零点设在入射光束的中心.

拉曼-奈斯型衍射发生在低声频和当声光相互作用长度不太大的情况下，产生的条件

$$l \ll \frac{\Lambda^2}{\lambda}. \tag{6.5.4}$$

由于相互作用长度短，因此没有多重衍射发生，入射光被衍射成为一系列不同级次的光，分别以不同角度出射

$$\sin\theta = \frac{m\lambda_0}{\Lambda}, \quad m = 0, \pm 1, \pm 2, \cdots, \tag{6.5.5}$$

这些级次光的强度与没有声光作用时的光强 I_0 之比为

$$I/I_0 = \begin{cases} [\mathrm{J}_m(\Delta\varphi')]^2/2, & |m| > 0, \\ [\mathrm{J}_0(\Delta\varphi')]^2, & m = 0, \end{cases} \tag{6.5.6}$$

其中 J_m 是 m 阶贝塞尔函数，$\Delta\varphi'$ 是 $\Delta\varphi$ 的最大值. 一般而言，拉曼-奈斯型调制器都采用零阶模或一阶模输出，衍射输出的光波振幅受到了超声功率的调制，可用此来做小信号的调制器.

调制度定义为

$$\eta_{\mathrm{RN}} = \frac{[I_0 - I(m=0)]}{I_0} = 1 - [\mathrm{J}_0(\Delta\varphi')]^2. \tag{6.5.7}$$

一般地，拉曼-奈斯型调制器的调制度比布拉格调制器的调制度小，而且由于拉曼-奈斯型调制器的衍射光以不同的角度分散在不同的地方，因此拉曼-奈斯型衍射不适合用来做开关，也很少在集成光学中应用.

2. 布拉格型调制

布拉格衍射基本结构如图 6.27 所示，它是指当入射光以一定的角度入射时，衍射光谱由相应于 $m=0$ 和 $m=1$ 的极值组成，负的 1 级和高级的衍射极值是不存在的，1 级的强度将为最大.

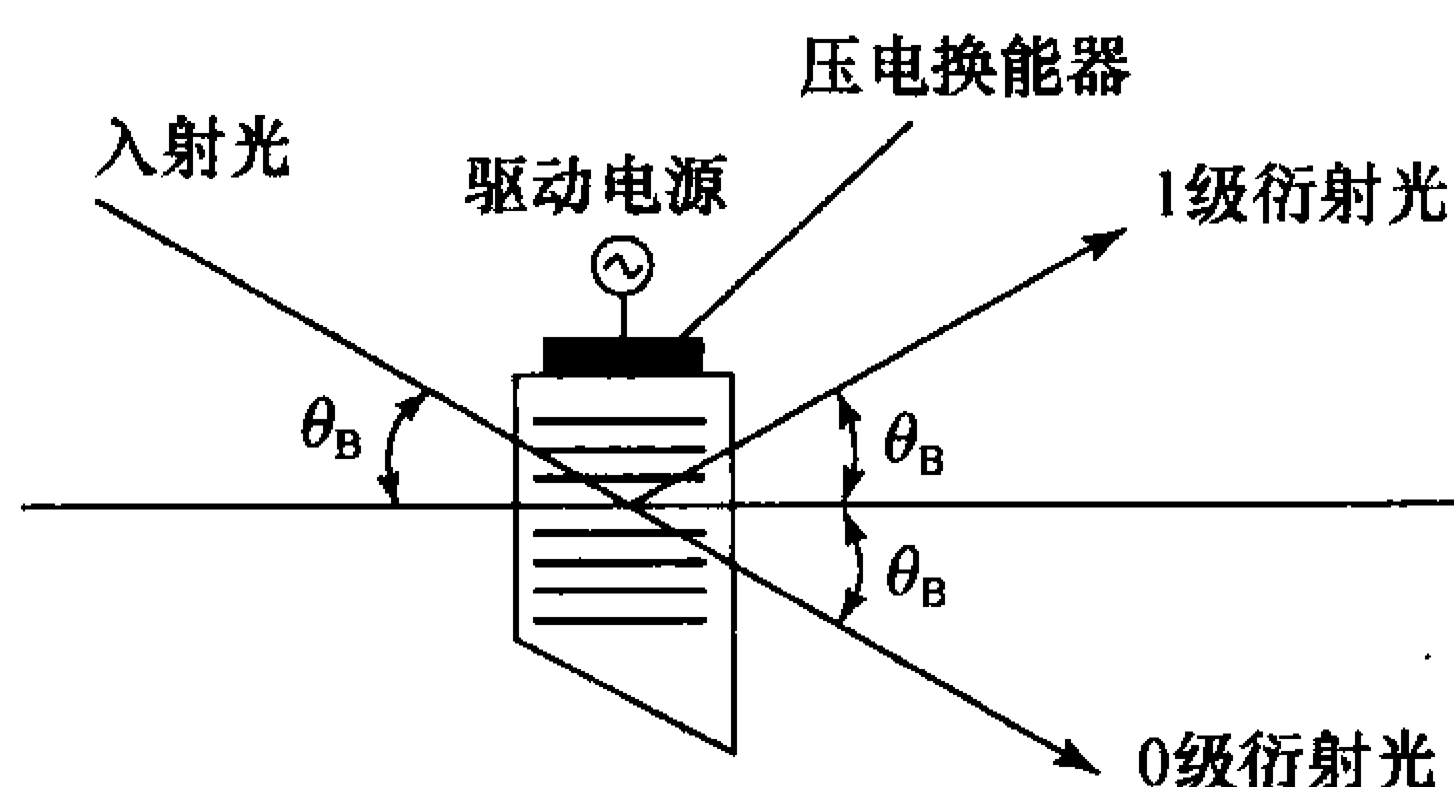


图 6.27　布拉格衍射原理图

布拉格衍射发生在声与光相互作用长度大和频率高的情况下,需同时满足入射角和声光作用长度的条件,即为

$$\sin\theta_B = \lambda/2\Lambda, \tag{6.5.8}$$

$$l \gg \frac{\Lambda^2}{\lambda}, \tag{6.5.9}$$

其中 θ_B 称为布拉格角,0 级和 1 级输出相对光强为

$$I_0 = \cos^2(\Delta\varphi/2), \tag{6.5.10}$$

$$I_1 = \sin^2(\Delta\varphi/2). \tag{6.5.11}$$

当 $\Delta\varphi=\pi$ 时,1 级衍射效率为 100%.

一般地,以 0 级光束为输出调制,调制度为

$$\eta_B = \frac{I_0 - I}{I_0} = \sin^2\left(\frac{\Delta\varphi'}{2}\right). \tag{6.5.12}$$

基于布拉格衍射型体调制器现在已经被广泛应用.集成光学中硅基波导声光调制器现阶段在发展,其工作方法与体调制器相似,而且还具有较小的驱动超声功率的优点,硅基声光调制器主要采用布拉格衍射实现.

6.5.2 硅基声光调制器结构参数及性能指标

1. 结构参数

声光调制器由声光介质、电声换能器、吸声装置及驱动源组成.传统声光调制的声光介质有熔融石英、重火石玻璃、钼酸铅晶体、氧化锌薄膜等,一般由具有高的声光效应的材料制作,在硅基中有采用铝砷化镓、石英[46]等.

换能器为由射频压电换能器组成的超声波发生器.压电换能器是声光器件的重要组成部分,它的作用是接受输入的电信号,通过逆压电效应将电信号转换成相应的声信号,并把声信号传入声光互作用介质中.硅基换能器一般都采用交叉式金属薄膜制成,这种方式能很好地将电能转换为声功率,从而引起介质弹性形变,使得折射率发生改变.其一般的单周期的普通方式的换能器结构如图 6.28(a)所示,使用精确的光刻技术制作的换能器能产生 1 GHz 的声表面波,如果想得到更高频率的声表面波就需要精度更高的电子束刻蚀.换能器的制作对材料要求也非常高,要求材料有高的机电耦合系数,产生强的压电效应,从而产生强的声表面波,如氧化锌(ZnO)薄膜.

驱动源用以产生调制电信号施加于电声换能器的两端电极上,驱动电声换能器将电功率转换成声功率.吸声器用以吸收已通过介质的声波,超声波呈行波状态.为使超声场呈驻波状态,吸声装置应换成反射型装置.

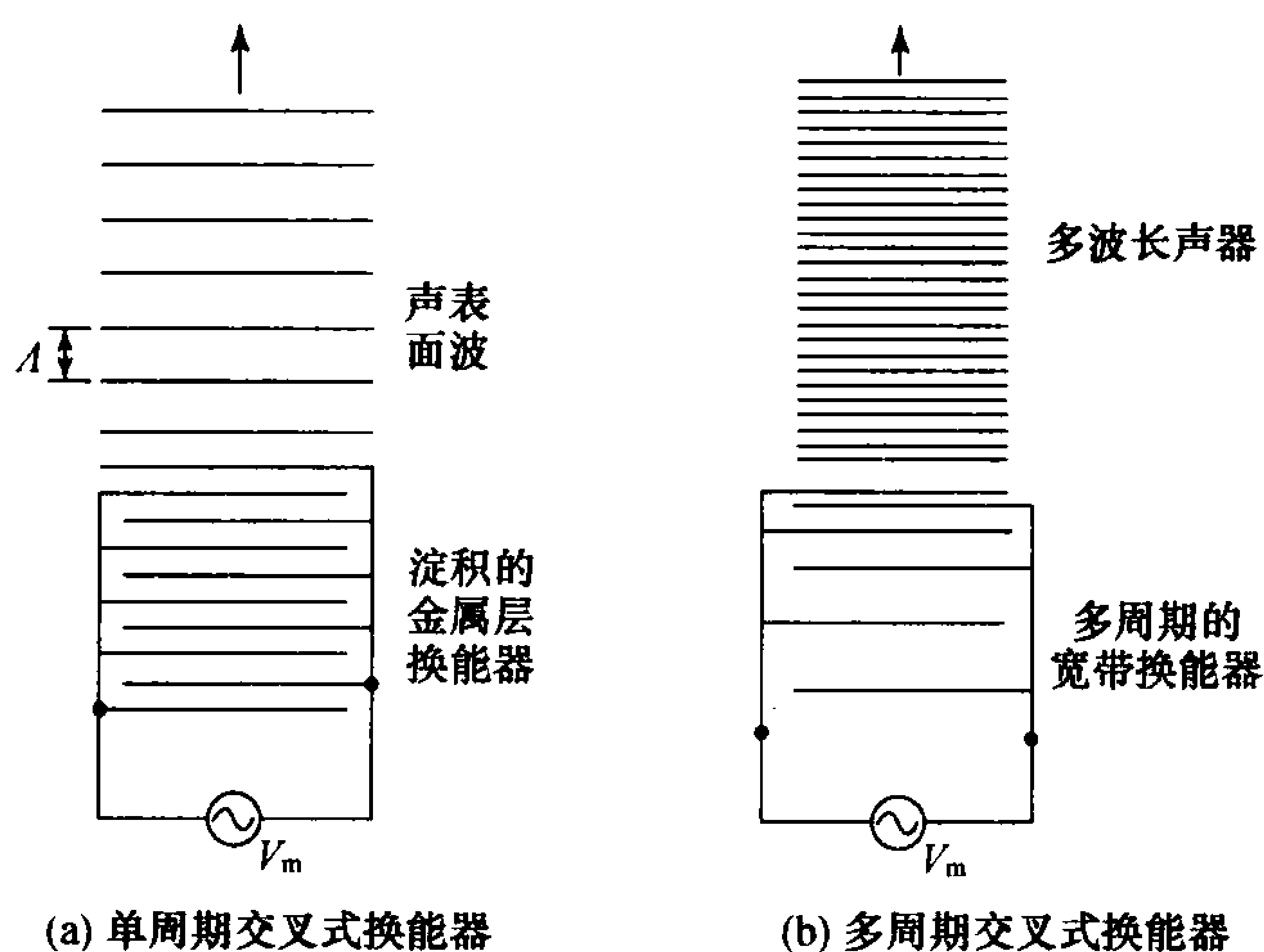


图 6.28 换能器结构

2. 性能指标

(1) 调制带宽.

声光调制带宽受超声功率的影响，也受声光相互作用的长度 l 的影响. 减少互作用长度 l，减少衍射效率，可以增加带宽，这也需要更高的超声功率. 如 Tsai 研究的 $LiNbO_3$ 基底 2 μm 厚光波导，声表面波频率为 700 MHz，l 为 3 mm，入射光波长为 6328 Å 时，调制 3 dB 带宽为 34 MHz；而将波导设计为 1 μm 厚，声表面波为 1 GHz，l 为 0.2 mm 时，调制带宽增加到 380 MHz[47].

如图 6.28(b)所示也可采用多周期交叉式换能器来提高调制带宽，多周期换能器在不同的地方产生不同的声波波长，从而增加全带宽[47].

(2) 调制度.

调制度类同于消光比，它是调制器性能好坏的一个指标，调制度越接近于 1 越好，输出的光最小值越接近零，调制深度就越深.

(3) 驱动功率.

驱动功率就是实现调制所需要的最小的电功率，它与换能器密切相关，换能器好，即使在小的驱动功率下，也能产生强的声表面波.

(4) 换能器效率.

换能器的一个重要性能指标换能器转换效率，换能器高效率转换损耗的功率小，所需功耗低，散发的热量小，系统性能稳定.

6.5.3 声光调制研究进展

第一个波导布拉格型声光调制器被报道是在 1970 年[48]，由 IBM 公司的 Kuhn 采用图 6.29 所示的方法实现的．基底是由压电效果好的石英形成，在石英表面溅射形成一层 0.8 μm 厚的玻璃薄膜波导，交叉型金属膜换能器在石英表面产生声表面波，由于机械接触，从而使得波导形成光栅，入射光波长为 6328 Å，声表面波频率为 191 MHz，波长 16 μm，实现调制度为 70％的调制．随后，Wille 和 Hamilton 采用同样的方法实现了布拉格型声光调制，调制度提高到 93％[46]．1973 年单片集成的声光调制器被研制出来，Chubachi 将声表面波直接驱动进入氧化锌波导中，采用布拉格衍射方式产生调制器，达到调制深度 90％以上，且驱动效率更高[49]．

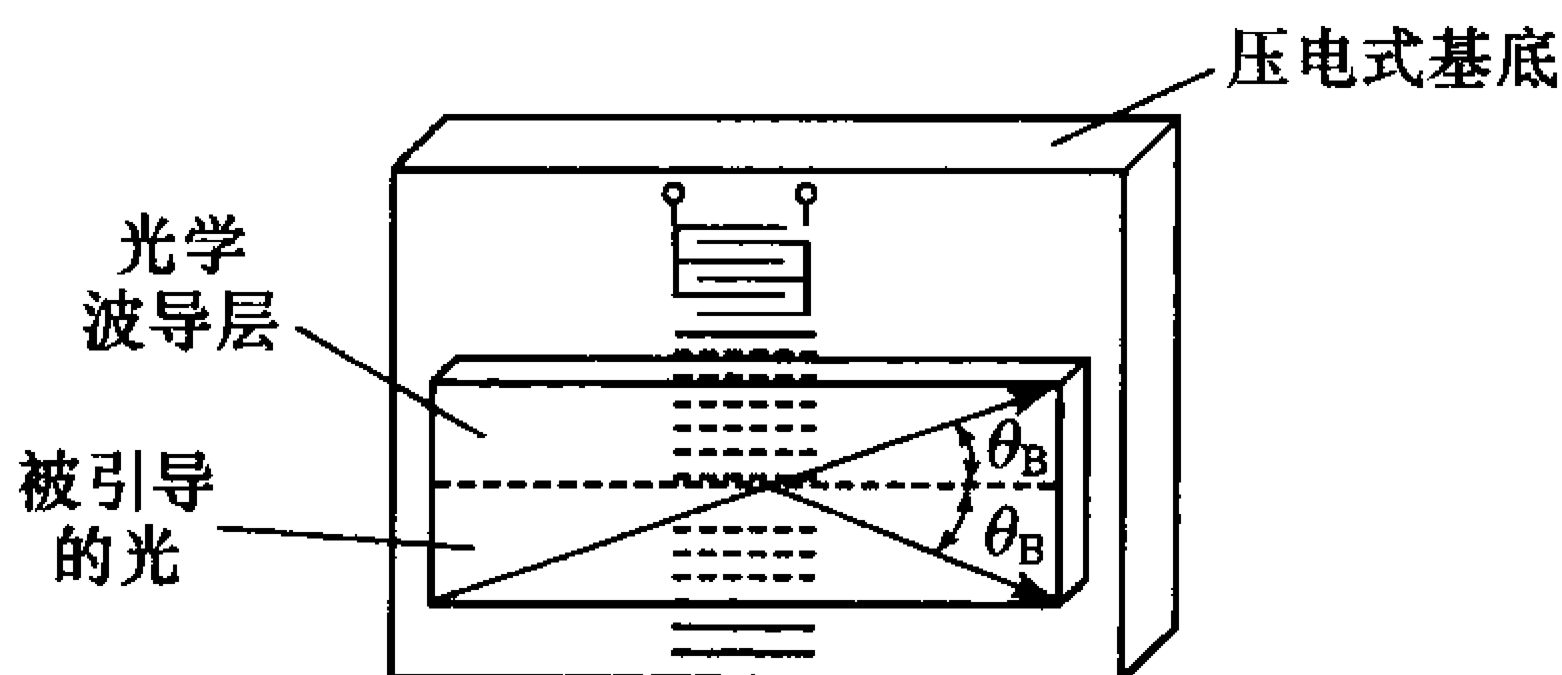


图 6.29 布拉格型波导声光调制器

1970 年，另一种利用声光互作用产生模式之间的耦合方法被提出，随后很多采用模式耦合方法的调制被研制，发现声光调制有实现在 GHz 进行调制的潜质品质[50]，通过声光效应产生光传输介质的折射率的改变可实现光强度调制[41]，1997 年，日本的 Gorecki 利用 MZI 结构在硅基上利用声表面波相位调制做出光集成干涉仪，其结构图如图 6.30 所示，采用在 2.5 μm 厚的 ZnO 薄膜上沉积一层铝来实现交叉式电极，通过 ZnO 薄膜换能器来产生声表面波，MZI 的一个臂的光受声表面波的影响，产生相位变化，从而在输出端产生干涉[51,52]．

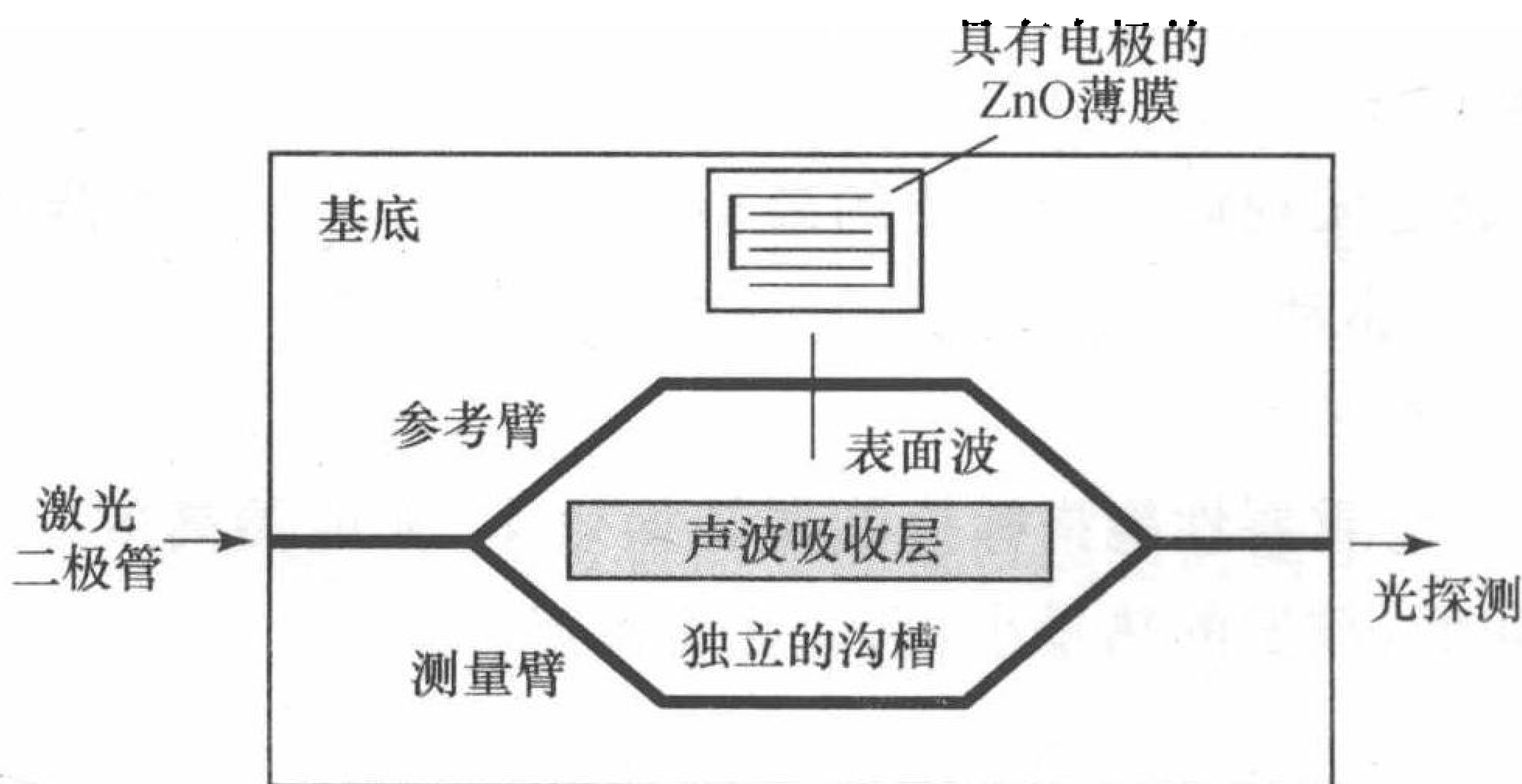


图 6.30 基于 MZI 的声光相位调制

德国 M. M. de Lima 于 2006 年在Ⅲ-Ⅴ族上利用声光效应实现了 GHz 量级的光强度操作，其结构如图 6.31 所示，在之前日本的基础上，做了两点改动，一是声表面波对 MZI 的两个臂都产生影响，MZI 两臂间距是声表面波半波长的奇数倍，由于声表面波是行波，输出的两束光在输出端能产生干涉调制，而且调制的频率能达到声表面波的两倍；另一个是采用了一个更集中化的交叉式换能器，能产生一个更窄更强的声表面波，提高功耗，减少了尺度[53]。

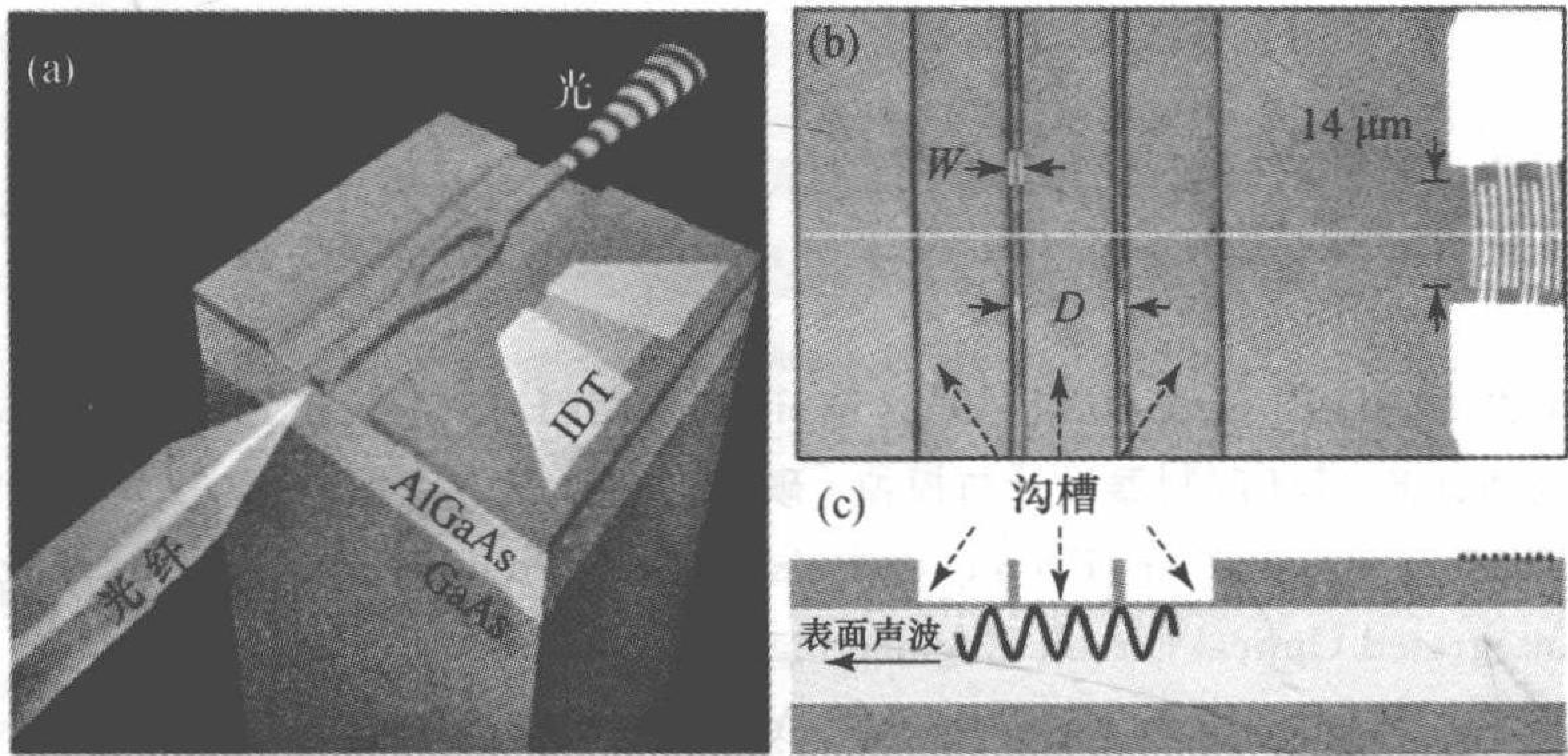


图 6.31　基于 MZI 的声光调制原理图

在硅基上真正实现第一个声光调制器是美国康奈尔大学的 S. Sridaran 和 S. A. Bhave 在 2010 年制作出来的，其结构如图 6.32 所示，它是由射频在左边环中产生超声，引起左边环的机械谐振，通过中间耦合束引起右边光谐振环的半径变化，从而引起波导透射光谐振波长的变化，实现光强调制的功能，在 1550 nm 光下能实现 288 MHz 的声光调制[54]。

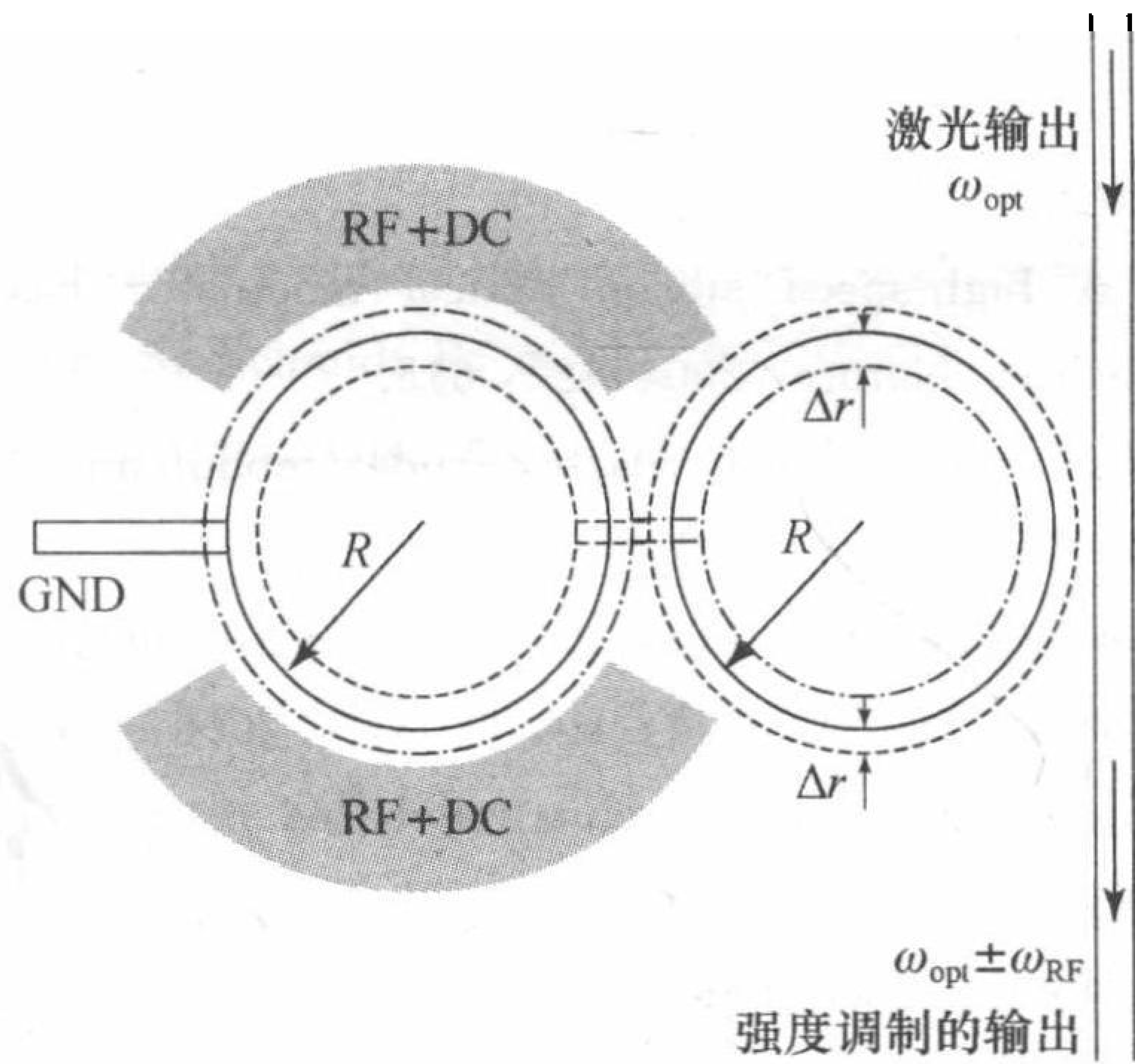


图 6.32　微环声光调制原理图

6.5.4 小结

该节通过声光效应介绍了传统的几种结构的声光调制结构和工作原理，并由此介绍硅基声表面波的声光调制结构成分及性能参数，主要以布拉格型声光调制为主，后又介绍了几种利用声光效应引起折射率改变从而实现调制功能的调制器，最后总结了国内外声光调制的应用及发展状况.

参考文献

[1] 徐介平. 声光器件的原理、设计和应用. 北京：科学出版社，1982.

[2] Pinnow D. Guide lines for the selection of acoustooptic materials. IEEE J. ,1970, QE-6: 223.

[3] Soref R A, et al. Electrooptical effects in silicon. IEEE J. Quant. Electr. 1987,23: 123.

[4] 毛岸. 高速硅基 MZI 调制器设计与模拟 [硕士论文]. 武汉：华中科技大学，2008.

[5] Soref R A, et al. Kramers-Kronig analysis of electro-optical switching in silicon. Proc. SPIE Integrated Optical Circuit Engineering Ⅳ, 1987.

[6] Reed G T, Mashanovich G, Gardes F Y, et al. Silicon optical modulators. Nature photonics,2010, 4: 518.

[7] Kuo Y H, et al. Strong quantum-confined Stark effect in germanium quantumwell structures on silicon. Nature, 2005,437: 1334.

[8] Rong Y, et al. Quantum-confined Stark effect in Ge/SiGe quantum wells on Si. IEEE J. Sel. Topic. Quant. Electr. , 2010,16: 85.

[9] Liao L, et al. 40 Gbit/s silicon optical modulator for high-speed applications. Electr. Lett. , 2007,43: 1196.

[10] Png C E, et al. Optical phase modulators for MHz and GHz modulation in silicon-on-insulator (SOI). J. Lightwave Technology, 2004,22: 1573.

[11] Preston K, et al. Deposited silicon high-speed integrated electro-optic modulator. Opt. Exp. , 2009,17: 5118.

[12] Liu A, et al. A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor. Nature, 2004,427: 615.

[13] Liao L, et al. High speed silicon Mach-Zehnder modulator. Opt. Exp. , 2005, 13: 3129.

[14] Basak J, et al. Developments in gigascale silicon optical modulators using free carrier dispersion Mechanisms. Adv. Opt. Tech. , 2008, V 2008: 678948.

[15] Liu A, et al. High-speed optical modulation based on carrier depletion in a silicon waveguide. Opt. Exp. , 2007,15: 660.

[16] Marris M D, et al. Optical modulation by carrier depletion in a silicon p-i-n diode. Opt. Exp. , 2006,14: 10838.

[17] Gardes F Y, et al. A sub-micron depletion-type photonic modulator in silicon on

insulator. Opt. Exp., 2005, 13: 8845.

[18] Gunn C. CMOS photonics for high-speed interconnects. IEEE Micro, 2006, 26: 58.

[19] Soljacic M, et al. Photonic-crystal slow-light enhancement of nonlinear phase sensitivity. J. Opt. Soc. Am. B: Opt. phys., 2002, 19: 2052.

[20] Gu L, et al. High speed silicon photonic crystal waveguide modulator for low voltage operation. Appl. Phys. Lett., 2007, 90: 071105.

[21] Tanabe T, et al. Low power and fast electro-optic silicon modulator with lateral p-i-n embedded photonic crystal nanocavity. Opt. Exp., 2009, 17: 22505.

[22] Chen X, et al. Capacitor-embedded 0. 54 pJ/bit silicon-slot photonic crystal waveguide modulator. Opt. Lett., 2009, 34: 602.

[23] Wei J, et al. Photonic crystal waveguide modulators for silicon photonics: Device physicsand some recent progress. Solid-State Electronics, 2007, 51: 1278.

[24] Li J, et al. Systematic design of flat band slow light in photonic crystal waveguides. Opt. Exp., 2008, 16: 6227.

[25] Hochberg M, et al. Terahertz all-optical modulation in a silicon-polymer hybrid system. Nature Materials, 2006, 5: 703.

[26] Kuo Y H, et al. High speed hybrid silicon evanescent electro absorption modulator. Opt. Exp., 2008, 16: 9936.

[27] Chen H W, et al. High speed hybrid silicon evanescent Mach-Zehnder modulator and switch. Opt. Exp., 2008, 16: 20571.

[28] Xu Q, Schmidt B, Pradhan S, et al. Micrometre-scale silicon electro-optic modulator. Nature, 2005, 435: 325.

[29] Manipatruni S, et al. Wide temperature range operation of micrometer-scale silicon electro-optic modulators. Opt. Lett., 2008, 33: 2185.

[30] Teed G T. Silicon Photonics — the state of the art. West Sussex: John Wiley & Sons Ltd., 2008.

[31] Vlasov Y, Green W M J, Xia F. High-throughput silicon nanophotonic wavelength-insensitive switch for on-chip optical networks. Nature Photonics, 2008, 2: 242.

[32] Xu Q F, Manipatruni S, Schmidt B, et al. 12. 5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators. Opt. Exp., 2007, 430.

[33] Png C E. Silicon-on-insulator phase modulators [Ph. D. thesis]. Guildford: Univ. Surrey, 2004.

[34] Manipatruni S, Preston K, Chen L, et al. Ultra-low voltage, ultra-small mode volume silicon microring modulator. Opt. Exp., 2010, 18: 18235.

[35] Gill D M, Rasras M, Tu K Y, et al. Internal Bandwidth Equalization in a CMOS-compatible Si-ring Modulator. IEEE Photo. Tech. Lett., 2009, 21: 200.

[36] Ye T, Zhou Y F, Yan C S, et al. Chirp-free optical modulation using a silicon push-pull coupling microring. Opt. Lett., 2009, 34: 785.

[37] Ye T, Cai X R. On power consumption of silicon-microring-based optical modulators. J. Lightwave Technology, 2010,28: 1615.

[38] Wang X L. SOI thermo-optic modulator with fast response. CPL, 2003,1(9).

[39] Tinkera M T, Lee J B. Thermo-optic photonic crystal light modulator. Appl. Phys. Lett., 2005,86(22): 1111.

[40] Geuzebroek D H. Wavelength tuning and switching of a thermo-optic microring resonator. Proc. ECIO, 2003,1: 395.

[41] Watts M R. Ultralow power silicon microdisk modulators and switches. Proc. 5th. IEEE. Proc. Conference on Group IV Photonics, 2008.

[42] 魏红振. 快速响应 MNO 马赫曾德尔热光调制器. 半导体学报,2002,23(5): 509.

[43] Watts M R. Adiabatic resonant microrings (ARMs) with directly integrated thermal microphotonics. Proc. Conference on Lasers and Electro-Optics/Conference on Quantum Electronics and Laser Science, 2009.

[44] Dong P. Wavelength-tunable silicon microring modulator. Opt. Exp., 2010, 18(11): 10941.

[45] Gan F. Maximizing the thermo-optic tuning range of silicon photonic structures. Proc. IEEE Conference on Photonics in Switching, 2007.

[46] Tsai C. Guided-wave acousto-optic Bragg modulators for wide-band integrated optic communications and signal processing. IEEE Trans. Circ. Syst. 1979: 26: 1072.

[47] Pavesi L, Lockwood D J. Silicon Photonics. Berlin: Springer, 2004.

[48] Kuhn L, Dakss M L, Heidrich P F, et al. Deflection of an optical guided wave by a surface acoustic wave. Appl. Phys. Lett., 1970,17: 265.

[49] Chubachi N, Kushibiki J, Kikuchi Y. Electron. Lett., 1973,9: 193.

[50] Shah M L. Fast acousto-optical waveguide modulators. Appl. Phys. Lett., 1973, 23: 75.

[51] Gorecki C. Silicon-based integrated interferometer with phase modulation driven by surface acoustic waves. Opt. Lett., 1997,22(23): 1784.

[52] Bonnotte E, Gorecki C. Guided-wave acoustooptic interaction with phase modulation in a ZnO thin-film transducer on an Si-based integrated Mach-Zehnder interferometer. J. Lightwave Technology, 1999,17(1).

[53] de Lima M M. Compact Mach-Zehnder acousto-optic modulator. Appl. Phys. Lett., 2006, 89(12): 1104.

[54] Sridaran S, Bhave S A. Silicon monolithic acousto-optic modulator. Proc. 23rd IEEE International Conference on Micro Electro Mechanical Systems, 2010.

第七章　硅基光电探测器

光电探测器是将入射光能量转化为电信号的光电子器件，它不像激光器那样，必须是直接带隙的材料，因此硅虽然是间接带隙材料，也可以制备探测器. 但由于硅的禁带宽度是 1.12 eV，对波段在 1.1 μm 以上的红外光是透明的，而硅基光电子学所用波段主要是近红外波段，单纯体硅不适合做硅基光电子学所需要的红外探测器，具有良好光响应特性的红外光探测器也因此成为目前研究的重点. 其中，锗在红外波段具有高的响应，同时其制备技术和 CMOS 工艺兼容，因而锗硅探测器成为目前最有前景的硅基红外光电探测器. 本章将首先介绍光电探测器的基本原理；然后分析硅基光电探测器的结构和特性；最后在此基础上着重介绍两种探测器：体硅光电探测器和锗硅光电探测器. 在锗硅光电探测器方面，主要介绍了 p-i-n，APD 和 MSM 三种波导型锗硅光电探测器的发展过程和最新的研究结果.

§7.1　光电探测器的基本原理

7.1.1　光电探测器对光信号的吸收

光电探测器对光信号的吸收，主要有三种方式：带间光吸收、自由载流子光吸收和双光子吸收.

1. 带间光吸收[1,2]

能量为 $h\nu$ 的光入射到直接带隙为 E_g 的半导体材料上，如果入射光的光子能量比半导体的禁带宽度 E_g 大，则位于价带的电子就可能吸收光子的能量而被激发至导带中，这一过程称之为带间光吸收.

在传播介质中，$x=0$ 处光强为 I_0 的光沿着 x 轴方向传播一段距离 x 后，其光强会因介质的光吸收而下降为 $I(x)$，它遵循如下规律：

$$I(x) = I_0 e^{-\alpha x}, \tag{7.1.1}$$

即光强会因光吸收而成指数下降. 式中，α 为吸收系数. 当光传输了一段距离 $d=1/\alpha$ 时，$I(d)=I_0/e$，即光强正好为原来的 $1/e$，d 为光波的穿透深度. 图 7.1 给出了实验测得的 Si，Ge 和 GaAs 的光吸收系数同光波波长和光穿透深度的关系[1]. 由图可见，光波波长越短，光吸收系数越大，光穿透深度越短，

实验表明 α 同波长平方根的倒数成正比关系. 由于硅和锗是间接带隙半导体,所以曲线缓慢到截止波长,而 GaAs 是直接带隙半导体,所以在截止波长曲线比较陡直.

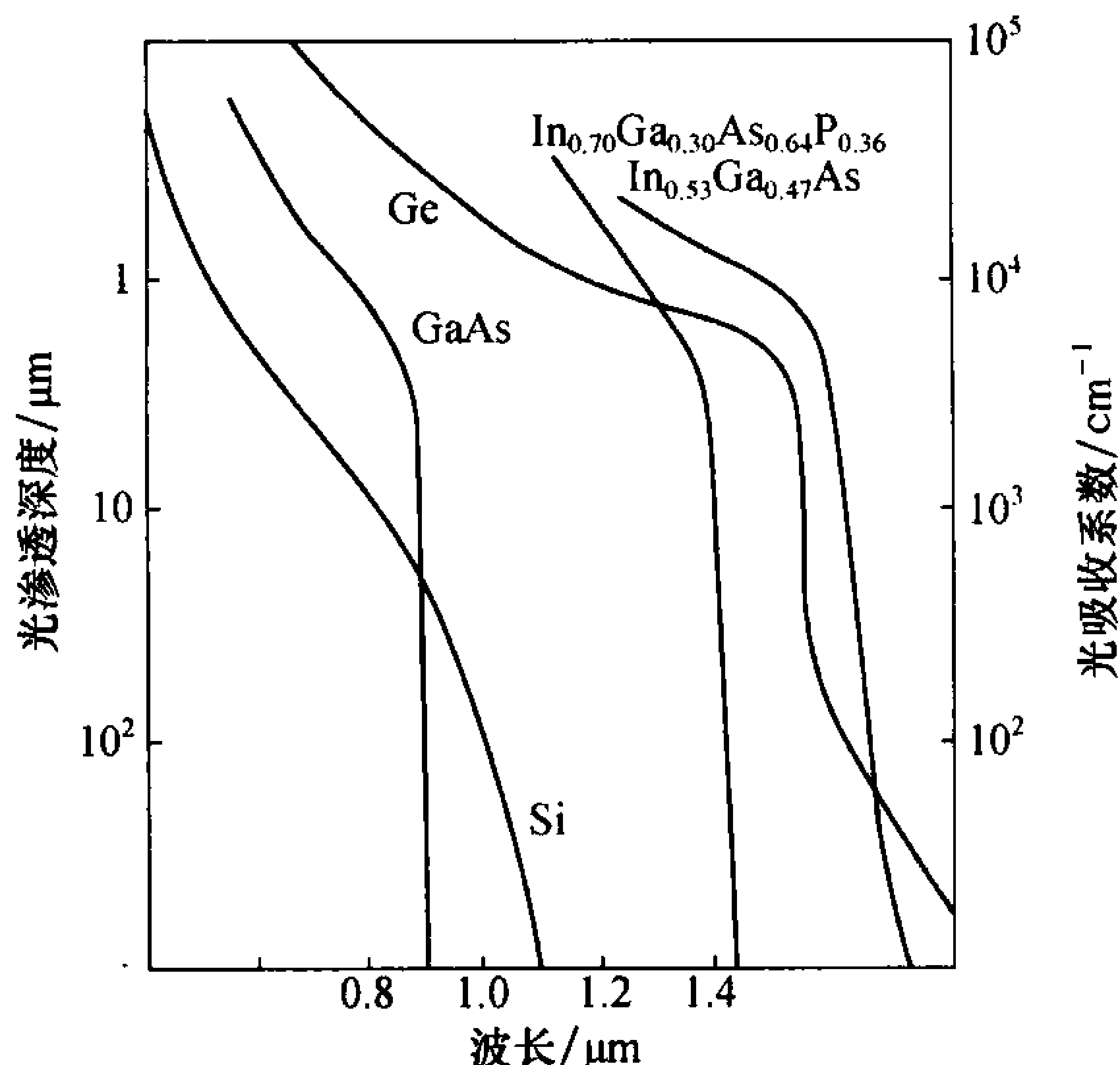


图 7.1 实验测得的 Si,Ge 和 GaAs 的光吸收系数同光波波长和光穿透深度的关系

上述讨论中,只考虑导带的形状对吸收系数的影响. 在实际的半导体中,既要考虑导带形状的影响,还要考虑价带形状、杂质能级等对光吸收的影响. 掺杂半导体中,对于施主能级 E_D 和受主能级 E_A,无论是 E_D-E_v(价带能级),E_c(导带能级)$-E_v$,还是 E_D-E_A 都小于带隙 E_g. 当电子吸收光子的能量实现电子由 E_A 到 E_c 或 E_V 到 E_D 的杂质-带边的跃迁、或 E_A 到 E_D 的杂质-杂质的跃迁,都会在价带或导带中产生自由载流子,使杂质能级电离. 这种杂质能级参与的过程,叫做杂质能级光吸收. 在间接带隙半导体中,光吸收过程还会有声子参与,因此需要采用二级微扰理论来处理光子吸收问题,这就变得复杂得多.

2. 自由载流子光吸收[1]

当入射光的光子能量 $h\nu$ 小于带隙 E_g 时,虽然不会造成带间跃迁的本征吸收,但在光场作用下,导带和价带中的载流子产生运动,也会出现光吸收现象,即自由载流子光吸收. 如果半导体中自由载流子浓度为 N,介电常数为 ε_0,电子的有效质量和迁移率分别为 m_e 和 μ_e,理论分析得出自由载流子的光吸收系数为

$$\alpha=\frac{e^5\lambda^2}{4\pi^2c^2n\varepsilon_0}\frac{N}{m_e\mu_e}, \tag{7.1.2}$$

其中 n 为介质的折射率. 可见自由载流子光吸收系数正比于自由载流子浓度，自由载流子浓度越高，则其光吸收越强. 同时，自由载流子吸收系数同波长的平方成正比，波长越长，吸收越大，表明自由载流子对长波吸收更为严重. 对于硅来讲[3]，$\alpha=10^{-18}\ n_{fc}\lambda^2$；当 $n_{fc}=10^{19}\ \mathrm{cm}^{-3}$，$\lambda=1.55\ \mu\mathrm{m}$ 时，$\alpha=24\ \mathrm{cm}^{-1}$.

3. 双光子吸收

通常原子从低能级跳跃到高能级，必须吸收一份相当于两个能级之差的能量. 如果这份能量由光辐射来提供，只有在光子的能量为两个能级之差时才会被原子所吸收. 但是在高功率的光束下，虽然一个光子的能量还达不到两个能级之差，但原子可以同时吸收两个光子达到一定的能量而完成一次跃迁，这就是双光子吸收.

7.1.2 光电探测器的基本原理[4]

光电探测器种类繁多，目前普遍使用的光电探测器是光电二极管(photodiode)和雪崩光电二极管(avalanche photodiode，APD). 它们是由半导体材料制作的. 由半导体材料制作的光电二极管，其核心部分是 p-n 结. p-n 结是 p 型半导体和 n 型半导体结合形成的. p 型半导体中空穴浓度高于电子浓度，而 n 型半导体中电子浓度高于空穴浓度. 由于扩散作用始终是浓度高的向浓度低的方向进行，所以当 p 型半导体和 n 型半导体结合在一起时，p 区的空穴将扩散到 n 区，而 n 区的电子将扩散到 p 区，使 p 区留下电离受主而带负电荷，而 n 区留下电离施主而带正电荷. 电荷堆积在 p-n 结两侧形成一自建电场，其方向由 n 指向 p. p 区的电子和 n 区的空穴在自建电场的作用下分别向 n 区和 p 区作漂移运动，同时 p-n 结的自建电场会阻止电子和空穴进一步向对方扩散而达到平衡，于是在 p-n 结区形成耗尽层.

为了提高光电二极管的响应速度，我们希望光生电子空穴对的产生尽量发生在耗尽层内. 因为在这一区域内一旦产生电子空穴对，电子和空穴立即被 p-n 结内强烈的自建电场分开而各自向相反方向作漂移运动，如图 7.2 所示. 由于自建电场很强，所以电子和空穴漂移运动的速度很快. 如果光生电子空穴对在耗尽层外部产生，由于耗尽层外不存在强烈的自建电场，电子和空穴只能靠扩散运动到达 p-n 结区，而扩散运动比漂移运动的速度低得多，所以将影响探测器的响应速度. 为了进一步提高响应速度，在实际使用时是将光电二极管反向偏置的，即将 n 接正，p 接负，外加电场方向与 p-n 结内自建电场方向一致. 这一外加电场使 p-n 结两侧的势垒差进一步加大，耗尽层宽度进一步加宽，允许更多的光生电子空穴对在高场强区产生，同时减小了二极管的结电容，从而进一步提高光电二极管的响应速度和灵敏度. 图 7.2 所示的是光电二极管的工作原理.

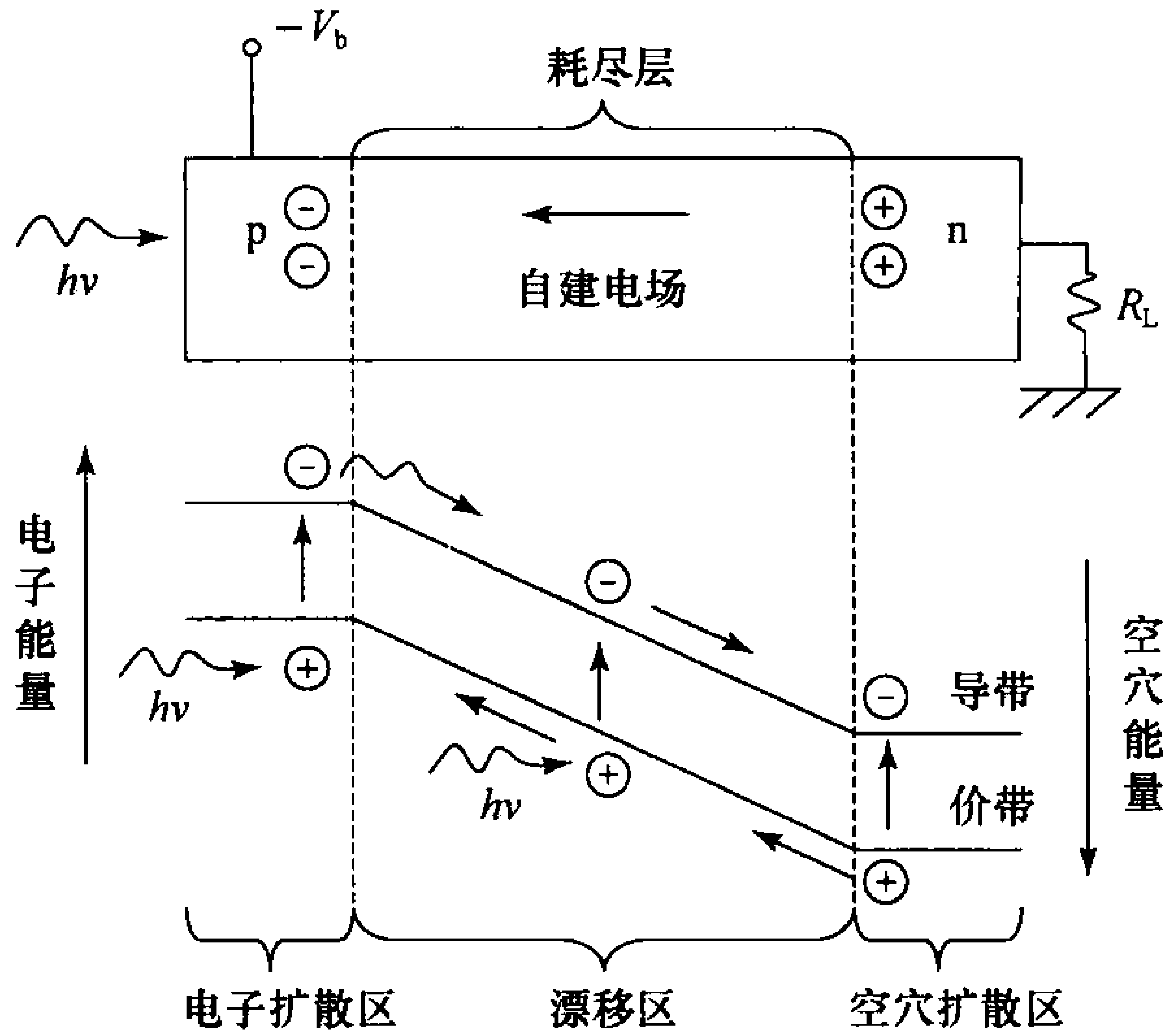


图 7.2 光电二极管工作原理

总之,半导体光电探测器是利用内光电效应进行光电探测,通过吸收光子产生电子空穴对,从而在外电路产生光电流.光电探测器的基本工作机理包括 3 个过程:① 材料在入射光照射下产生光生载流子;② 载流子输运或在电流增益机制下的倍增;③ 光电流与外围电路之间的相互作用并输出电信号.

§7.2 光电探测器的特性和结构

7.2.1 光电探测器的性能参数[5]

1. 量子效率

量子效率可分为内量子效率和外量子效率,它们是半导体光电探测器最重要的特性参数.内量子效率 $\eta_{内}$ 定义为吸收一个入射光子能够产生的电子-空穴对的个数,即

$$\eta_{内} = \frac{产生的电子-空穴对的个数}{入射光子数}.$$

由于 η 与材料的吸收系数 α,以及吸收层的厚度 W 相关,因而可表示为

$$\eta_{内} = 1 - e^{-\alpha(\lambda)W},$$

式中 $\alpha(\lambda)$ 是对应波长 λ 的吸收系数.由上式可见材料的吸收系数越大,或者吸收层越厚,光电探测器的量子效率就越高.在实际的光电探测器中,光不可能直接由材料表面达到吸收区,而是要经过一定的厚度的重掺杂接触区,在这个区域内

会造成一部分光子损耗，同时在光电探测器表面的反射作用也会损失部分入射光. 基于这些因素，定义外量子效率为

$$\eta_{外} = (1-R_f)e^{-\alpha(\lambda)d}\eta_{内} = (1-R_f)e^{-\alpha(\lambda)d}(1-e^{-\alpha(\lambda)w}),$$

式中 d 为前端接触层厚度，R_f 为光电探测器表面的反射率. 入射的一部分光线会由于界面的折射率不同而造成反射，反射率与界面的折射率 n_{sc}及吸收层消光系数 κ 有关. R_f 可以表示成如下形式

$$R_f = \frac{(1-n_{sc})^2+\kappa^2}{(1+n_{sc})^2+\kappa^2},$$

消光系数 κ 与吸收系数 α 的关系为

$$\alpha = \frac{4\pi\kappa}{\lambda_0}.$$

为了减小端面反射以提高外量子效率，可以在入射界面涂一层抗反射膜(anti-reflection coating, ARC)，抗反射膜厚度 d_{ARC}与波长和折射率有关，即

$$d_{ARC} = \frac{\lambda_0}{4n_{ARC}} = \frac{\lambda_0}{4\sqrt{n_s n_{sc}}}.$$

对于硅光电探测器，常用的抗反射膜材料为 SiO_2 和 Si_3N_4.

2. 响应度

入射到吸收区的光子产生的光生载流子在耗尽区内建电场的作用下，向探测器的两极漂移运动，并在输出端形成光电流. 通常在实际工作中，用单位入射光功率 P 与产生光电流 I_P 的关系，即响应度 R 来表示：

$$R = \frac{I_p}{P}. \tag{7.2.1}$$

根据对量子效率的定义，可得出

$$\eta = \frac{I_p/q}{P/h\nu}, \tag{7.2.2}$$

将式(7.2.2)代入式(7.2.1)中，可得

$$R = \eta\frac{q}{h\nu}, \tag{7.2.3}$$

其中 q 为电子电荷. 一般在实际探测器中采用表面镀抗反射涂层的办法来提高量子效率和响应度，通常能取得较好的效果.

3. 灵敏度

光电探测器的灵敏度是在系统一定的传输带宽和传输速率下所能检测到的最小光信号. 灵敏度也常称响应度，它是光电探测器光电转换特性、光电转换的光谱特性以及频率特性的量度，常用的单位是 dBm. 灵敏度又分为积分灵敏度 R、光谱灵敏度 R_λ 和频率灵敏度 R_f 三种[6].

4. 响应速度

光电二极管的响应速度是由探测信号的上升时间或下降时间来衡量的，通常

取两者之间较大的值,响应速度是光电二极管的一个重要参数.在光纤通信中要能够检测高频调制的光信号,从而提高响应速度和信噪比,降低系统的误码率.

5. 响应时间

当一束光入射到半导体的光电探测器上,入射光通过表面层进入半导体内,产生光生载流子,自由电子-空穴对,电子和空穴会在电场的作用下运动,分别朝相反的电极流去.只有电子空穴到达电极并形成光电流之后才能在外电路上探测出来.这整个过程需要一定的时间.由入射光转变为光电流所需的时间就是响应时间.

影响光电探测器响应时间的因素主要有:

① 耗尽区外载流子扩散时间 t_{diff}.载流子扩散运动较慢,而且大部分产生在耗尽区外的载流子寿命较短,很快就会复合.只有在耗尽区附近的部分载流子能够扩散进入耗尽区,并在电场作用下才能形成光电流.设载流子扩散系数为 D_c,则扩散距离 d 所需要的时间是

$$t_{\text{diff}} = \frac{d^2}{2D_c}. \tag{7.2.4}$$

② 耗尽区内载流子的渡越时间 t_{drift}.当耗尽区内的电场达到饱和时,载流子以最大漂移速度 V_d 运动,设定耗尽区宽度 W,则渡越时间为

$$t_{\text{drift}} = \frac{W}{V_d}. \tag{7.2.5}$$

W 的大小取决于施主和受主杂质的浓度,V_d 的大小取决于所加的偏置电压,并与材料有一定的关系.

③ 光电二极管的 RC 时间常数 t_{RC}.

$$t_{\text{RC}} = (R_L + R_s)C_p, \tag{7.2.6}$$

其中 R_L 为外部负载电阻,R_s 为内部串联电阻,C_p 是寄生电容.加大探测器本征层的宽度对提高量子效率和减小探测器寄生电容是一致的,而增大宽度意味着载流子在吸收区内的渡越时间会增加.减小探测器的面积可以有效减少结电容和暗电流,但是小面积的探测器给光纤的有效耦合带来了难度.因此,为了优化响应速度,需要选择适当的吸收层厚度和探测器面积.

6. 暗电流

在理想情况下,当没有光照射时,光电探测器应无光电流输出.但是实际上由于热激励(主要是由 p-n 结内热效应产生的电子空穴对形成)、宇宙射线及放射性物质的激励,在无光的情况下,光电探测器仍有电流输出,这种电流称为暗电流.它主要是由耗尽层中载流子的产生-复合电流和耗尽层边界的少数载流子扩散电流,以及表面漏电流构成.暗电流的大小与偏压和光电二极管的结面积有关,当偏压增大时,暗电流增大;与光电二极管的结面积成正比,故常用单位面积

上的暗电流(暗电流密度)来衡量. 一般取偏压是 0.9 V 时,对应的电流值为器件的电流值. 除了与偏置电压有关外,暗电流还随器件温度的增加而增加,从室温到 70℃,暗电流将增加一个数量级. 要减小器件的暗电流,首先应选择好的单晶材料,并严格工艺规程;其次要选定良好的表面钝化层,隔绝周围气体对器件的污染,在异质结光电二极管的设计中,有意将高场区移到宽禁带材料去,也是减小暗电流的有效措施.

7. 噪声

光电探测器的主要噪声源有散粒噪声、产生-复合噪声、光子噪声、热噪声和 $1/f$ 噪声. 散粒噪声是在无光照下,由于热激发作用而随机产生的电子所造成的起伏. 对光电导探测器,载流子热激发产生电子-空穴对,当电子和空穴运动时,也存在严重的复合过程,而复合过程本身也是随机的. 因此,不仅有载流子产生的起伏,而且还有载流子复合的起伏,这样就使起伏加倍. 虽然其本质也是散粒噪声,但为强调产生和复合两个因素,取名为产生-复合散粒噪声. 由于光激发的载流子也是随机的,也要产生起伏噪声,因为强调光子起伏,故称为光子噪声. 它是探测器的极限噪声,不论是信号光还是背景光,都要伴随着光子噪声,而且光功率越大,光子噪声越大. 热噪声是指: 由于光电探测器本质上可用一个电流源来等价,这就意味着探测器有一个等效电阻 R,因此,探测器的噪声可以用电阻器 R 两端随机起伏的电压来说明,这个起伏电压是由电阻中自由电子的随机热运动引起的. $1/f$ 噪声主要出现在大约 1 kHz 以下的低频频域,而且与光辐射的调制频率 f 成反比,故称为 $1/f$ 噪声或低频噪声. 实验发现,探测器表面的工艺状态(缺陷或不均匀等)对这种噪声的影响很大,所以有时也称为表面噪声或过剩噪声.

7.2.2　p-i-n 光电探测器

为了改善和提高光电二极管的性能,通常还在 p 区和 n 区之间形成一个本征 i 区(intrinsic region),构成所谓 p-i-n 光电二极管.

当在 p-n 光电二极管的 p-n 结间插入一层非掺杂或轻掺杂的半导体材料,就可以增大耗尽区宽度来减小载流子的扩散运动,提高响应速度. 由于 p-n 结中间插入的半导体材料近似为本征半导体,因此称为 p-i-n 光电二极管. 图 7.3 所示为 p-i-n 管的器件结构和反偏工作时的电场分布. 中间层本征材料具有高阻抗性质,使大部分电压加于其上,因而中间层存在一个较强的电场. 实质上耗尽区可扩展到 i 区之外,其宽度 W 可在制造过程中通过控制中间层厚度来调节. 由于 p-i-n 管中大部分入射光在 i 区被吸收,在这里产生的光生电子-空穴对将立即被电场分离,并快速漂移运动. 它与 p-n 光电二极管主要区别在于光电流的漂移分量相对于扩散分量占支配地位. 耗尽层的加宽也明显地减小了结电容 C_j,从而使电路的时间常数减小. 由于在光谱响应的长波区硅材料的吸收系数明显减

小,因此耗尽层的加宽还有利于对长波区光辐射的吸收,从而提供了较大的灵敏体积,有利于量子效率的改善. p-i-n 光电探测器的上述优点使它在光通信、光雷达及其他快速光电自动控制领域得到了非常广泛的应用.

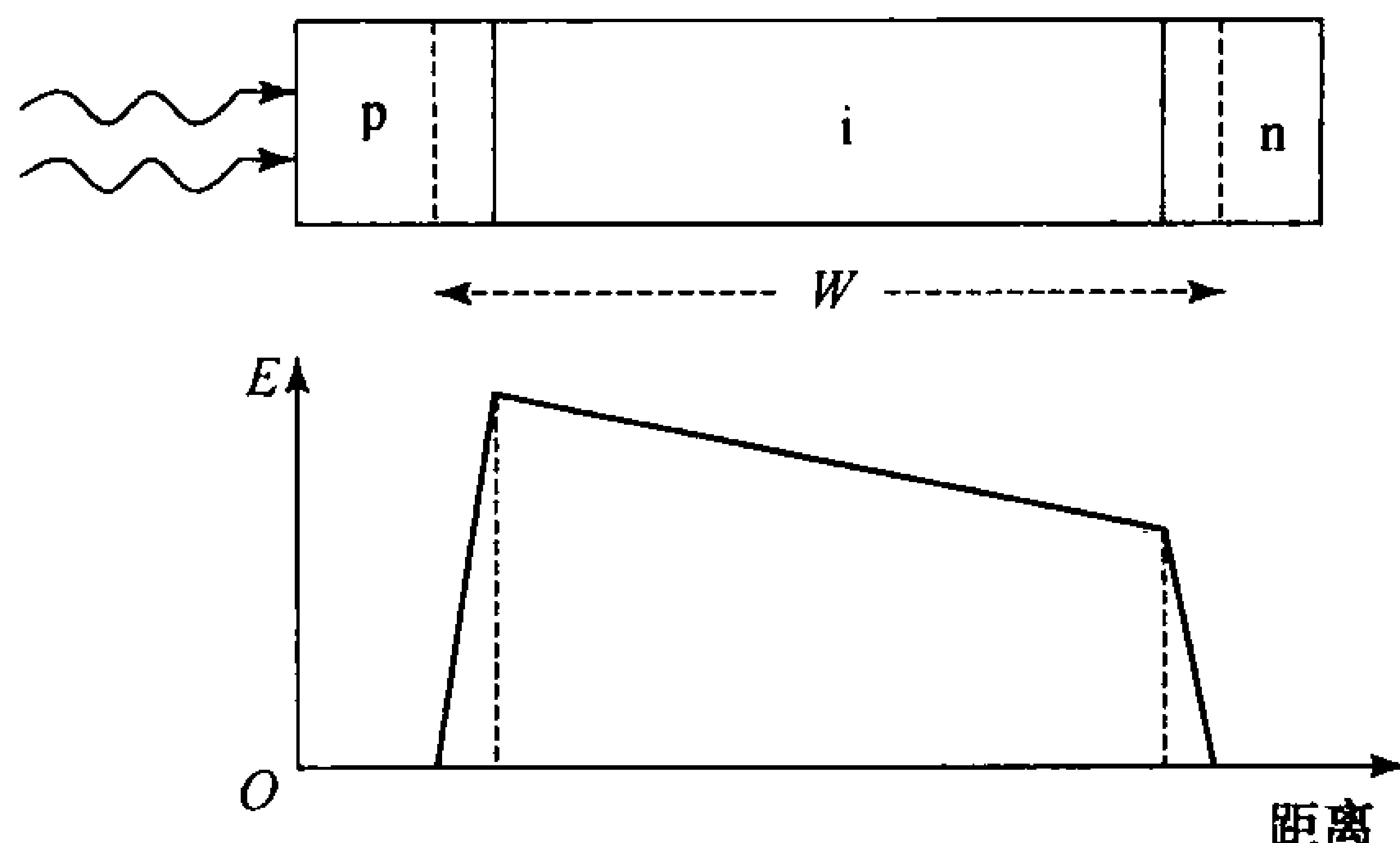


图 7.3 p-i-n 光电二极管及在反偏时各层的场分布

另外,虽然耗尽层宽度 W 增大可提高响应度,但是过大的 W 将导致响应速度变慢. 因此,耗尽区宽度要考虑灵敏度和带宽这两个指标合理折中选取.

7.2.3 雪崩光电探测器[7]

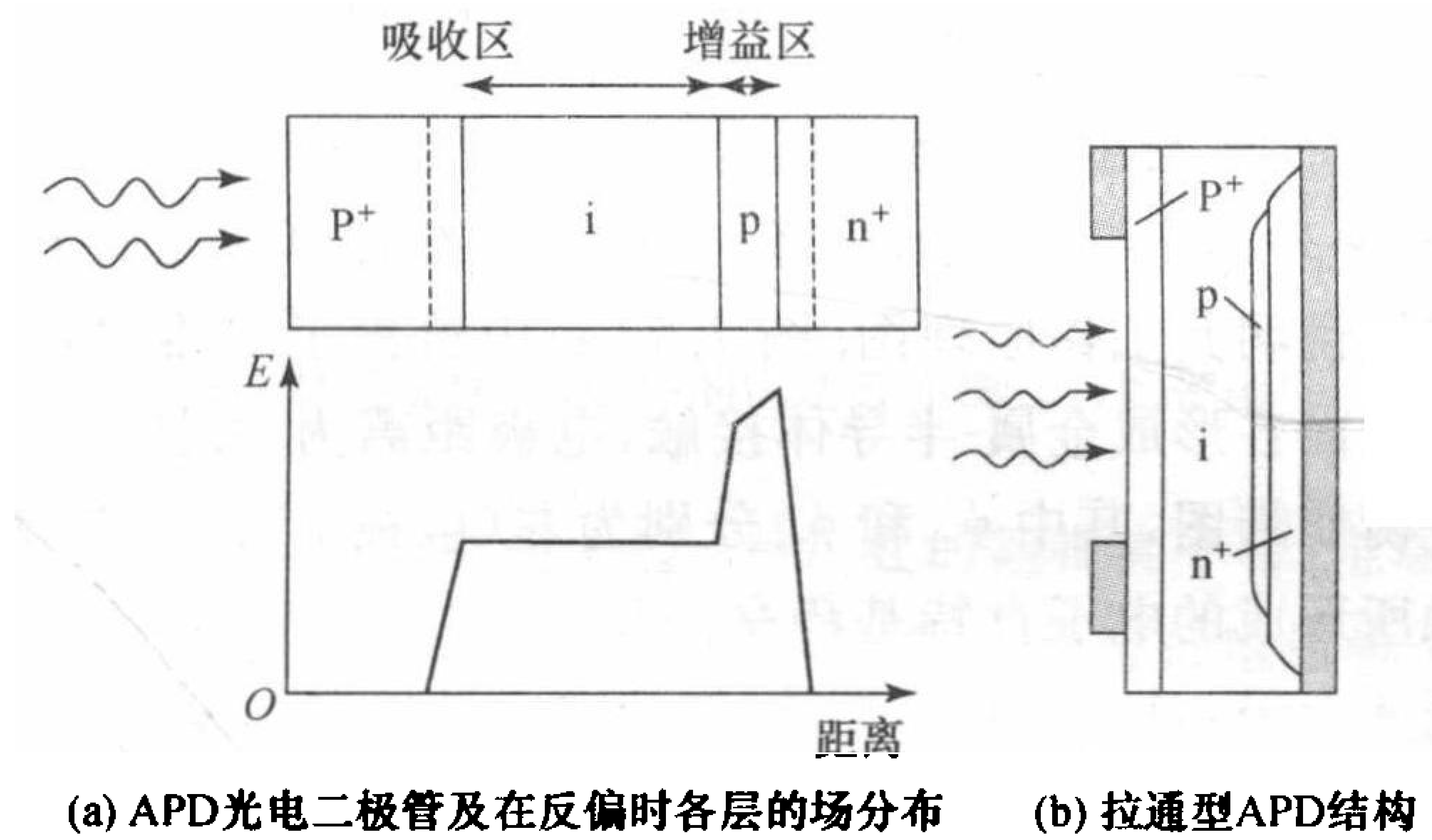
雪崩光电探测器是借助强电场作用产生载流子倍增效应(雪崩倍增效应)的一种高速光电器件. 一般来说,硅和锗雪崩光电探测器的电流增益可达 $10^2 \sim 10^3$,因此这种光电探测器的灵敏度很高且响应速度快,响应时间只有纳秒量级,相应的响应频率可达 100 GHz,噪声等效功率为 10^{-15} W. 它广泛应用于光纤通信、弱信号检测、激光测距、星球定向等领域. 任何探测器都存在一个保证正常工作的最小电流,也就存在一个最小入射功率 $P_{in} = I_p/R$,其中 R 为响应度. 探测器的响应度越高,其最小探测功率就越低. p-i-n 管的响应度在 $\eta = 1$ 时,达到最大值 $R = q/h\nu$. 而 APD 与光电倍增管类似存在内部电流增益,因而具有更高的响应度,能够探测更小功率的光信号. 它通常应用在光功率比较小的场合中.

APD 能够提供电流增益的物理机制在于碰撞电离. APD 在结构设计上就使其能承受高反向偏压,因而在 p-n 结内部形成一高电场区($\approx 3 \times 10^5$ V/cm). 一次光生的电子空穴对经过高电场区时会被加速,从而获得了足够的能量. 它们在高速运动过程中与晶格碰撞,使晶体中的原子发生电离,从而激发出新的电子空穴对,这个过程称为碰撞电离. 通过碰撞电离产生的电子空穴对称为二次电子空穴对. 新产生的电子和空穴在高电场区又被加速,又可能与晶格发生碰撞,使原子发生电离,产生新的电子空穴对. 这样多次碰撞电离的结果,使载流子迅速增加,反向电流迅速加大,形成雪崩倍增效应. APD 就是利用这种雪崩倍增效应

使一次光生电流获得放大.倍增率取决于两个参数：电子和空穴的碰撞电离系数 α_e 和 α_h，其值决定于半导体材料和电场强度.倍增系数为

$$M = \frac{1-\kappa_A}{\exp[-(1-\kappa_A)\alpha_e d]-\kappa_A}, \tag{7.2.7}$$

其中 $\kappa_A=\alpha_h/\alpha_e$.可见 APD 的电流增益与碰撞电离系数之比密切相关.通过研究可知，κ 越小，APD 的性能越高.APD 与 p-i-n 光电二极管的不同主要表现在增加了一个附加层，以实现碰撞电离产生二次电子-空穴对.图 7.4 所示为 APD 的结构，并给出了不同层中的电场分布.在反偏时夹在 i 层和 n 层间的 p 层中存在高电场，该层称为倍增区或增益区.大部分光子仍在耗尽层(i 层)中吸收，并产生一次电子-空穴对.耗尽层产生的电子空穴对经过增益区产生二次电子空穴对即可以获得增益.

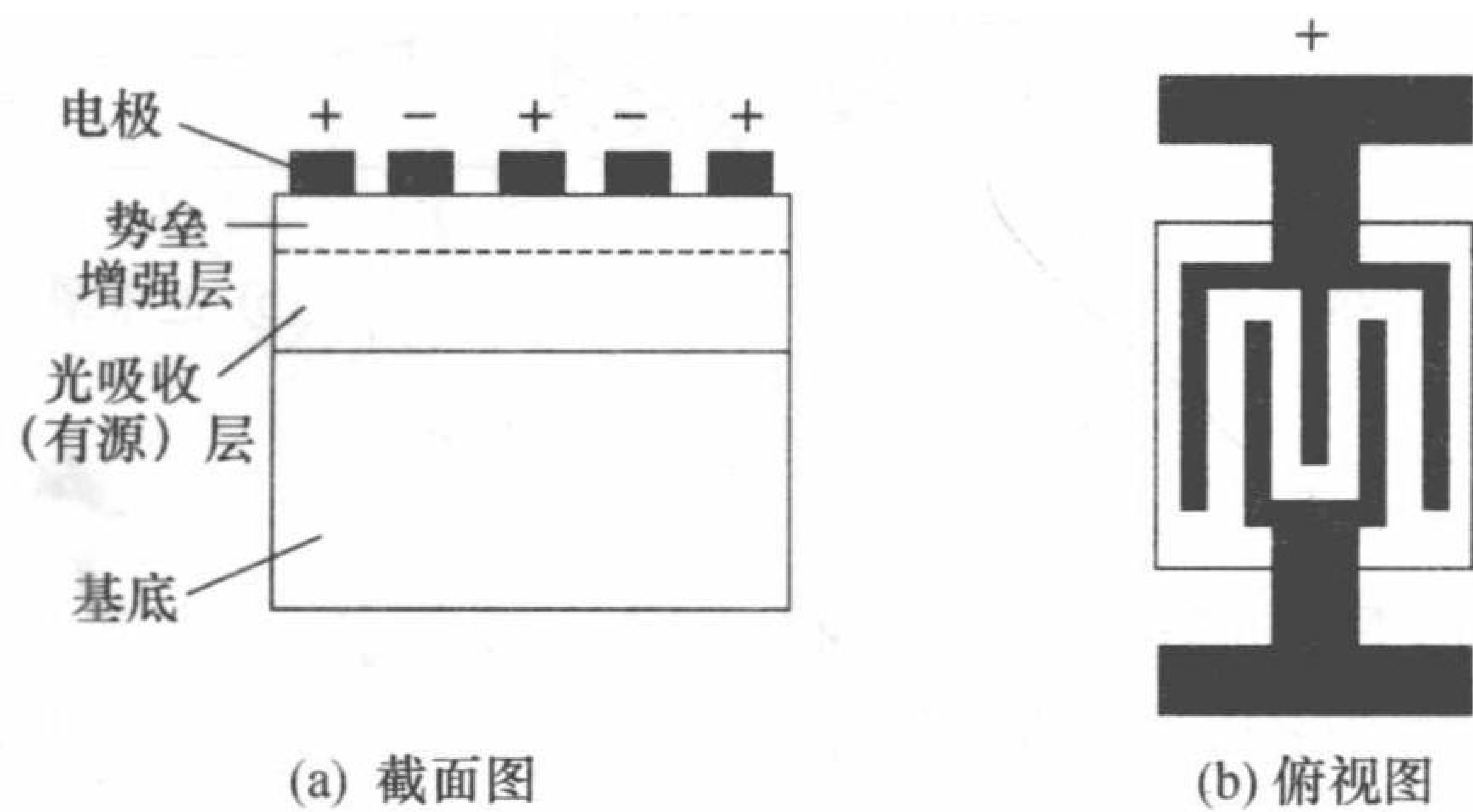


(a) APD光电二极管及在反偏时各层的场分布　　(b) 拉通型APD结构

图 7.4　APD 光电二极管的工作原理

7.2.4　金属-半导体-金属光电探测器[7]

金属-半导体-金属(metal-semiconductor-metal，MSM)结构的光电探测器本质上是一个背对背串联的两支金属半导体接触二极管，其典型结构如图 7.5 所示.



(a) 截面图　　(b) 俯视图

图 7.5　MSM 光电探测器器件结构

为了提高金属电极与有源层之间的肖特基势垒高度，有源层上面还有一层薄的势垒增强层；然后是用剥离技术形成的交叉指状金属电极；最后是在电极上面沉积的绝缘薄膜起钝化保护和抗反射作用[参见图 7.5(a)]. 图 7.5(b)为其俯视图可以看出交叉指状的金属电极形状. 当外加偏压时一个结为正偏置，另一个结为反偏置. 器件内电位分布如图 7.6 所示.

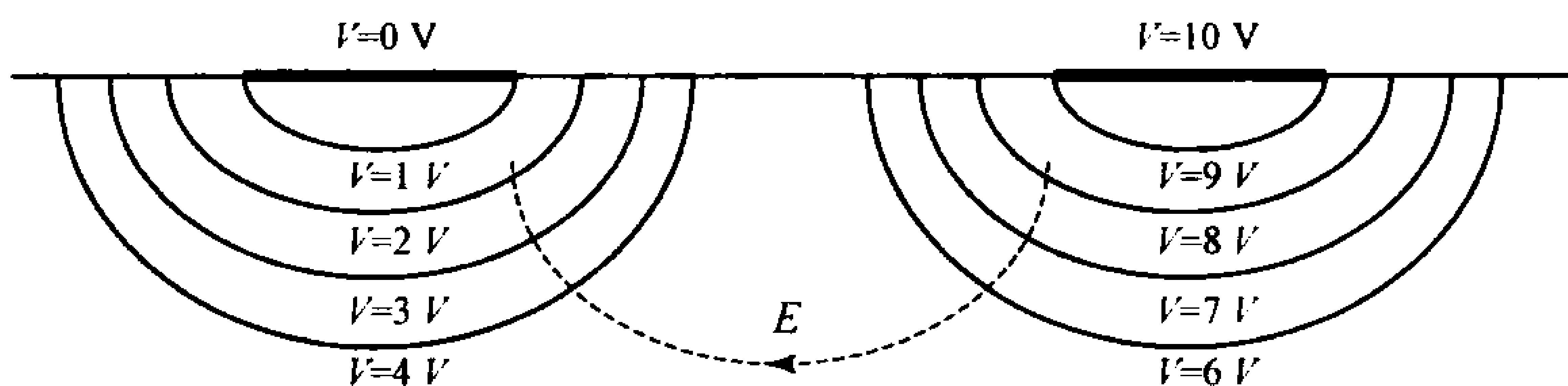


图 7.6 MSM 探测器电位分布图

施敏等在 1971 年给出了 MSM 结构的经典理论表述. 他们认为对于实际的应用结构，在势垒外的半导体上的压降一般很小，在计算时是将其忽略的. 下面用一维模型定性说明其工作原理图. 图 7.7(a)中给出了一维器件结构图，均匀掺杂半导体的两面各形成金属-半导体接触，电极距离为 L. 图 7.7(b)为外加偏压为零时的平衡能带图，其中 ϕ_{n1} 和 ϕ_{n2} 分别为左(电极 1)、右(电极 2)两面金属和半导体接触所形成的电子肖特基势垒高度，而 V_{D1} 和 V_{D2} 分别为其内建势. 对于同样金属则 $\phi_{n1}=\phi_{n2}$，$V_{D1}=V_{D2}$，而 ϕ_p 表示空穴的势垒高度. 当外加电压时(如右方为正，左方为负)，这一对背对背的二极管中，1 结为反偏置，2 结为正偏置. 其耗尽层宽度分别为 W_1 和 W_2. 随着外加电压的增加，反偏置的耗尽层宽度 W_1 增大，而正偏置的耗尽层 W_2 减小，但是其总的耗尽层宽度逐渐增加.

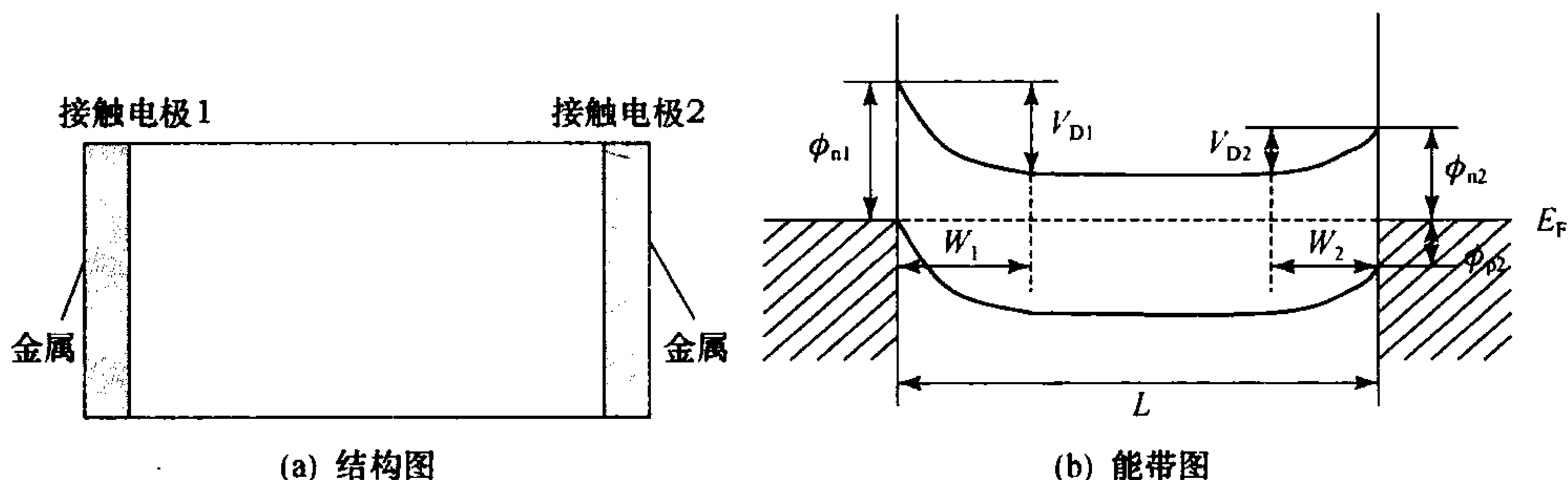


图 7.7 MSM-PD 一维器件

当使两耗尽层相接触时，相应的这一电压称为穿通电压 V_{RT}. 这时的电场分布和能带图如图 7.8(a)，(b)所示. 而且 $W_1+W_2=L$，在 x_0 点电场为 0，其左方电场为负方向，右方电场为正方向，这时仅有很小电流.

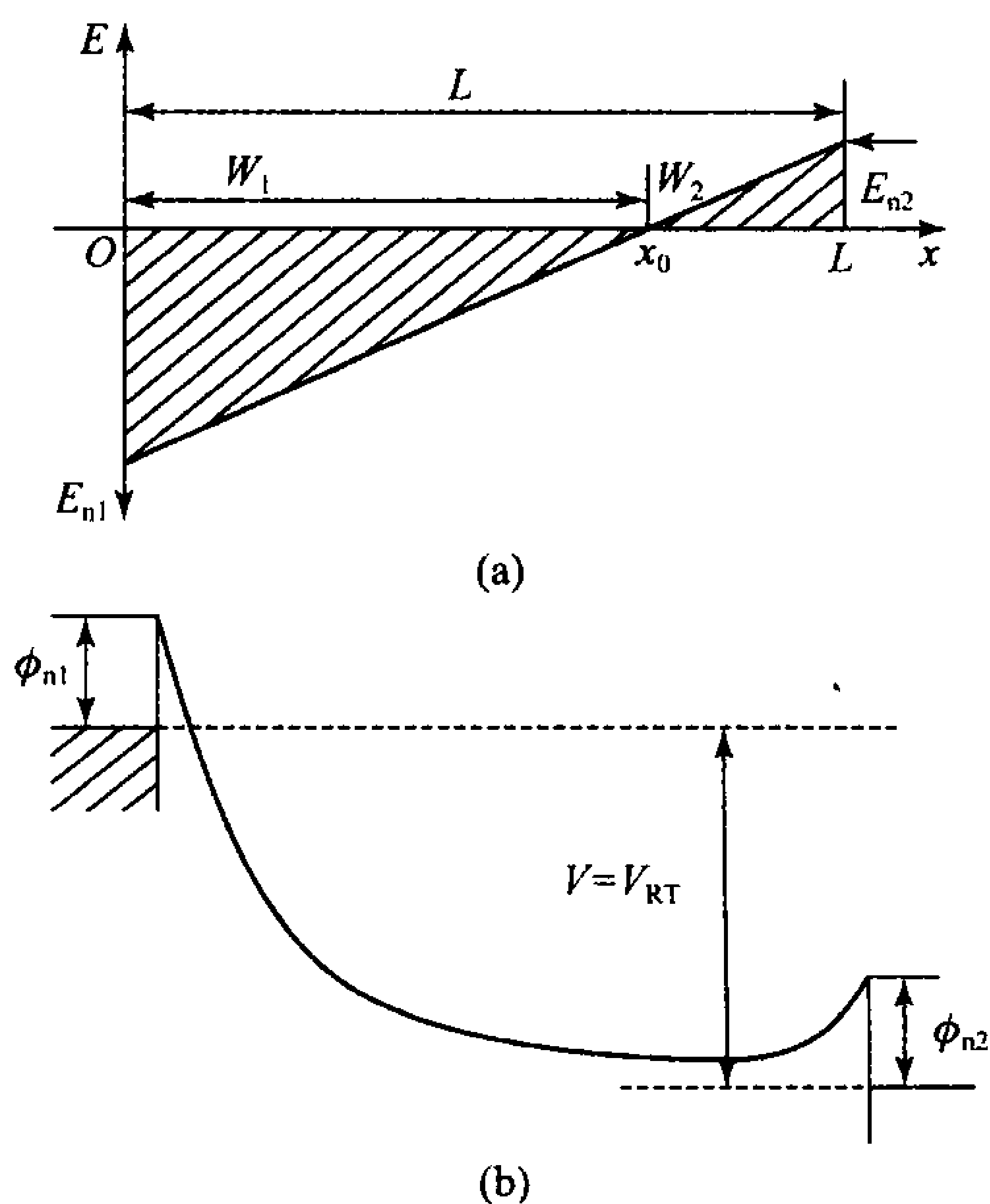


图 7.8　穿通条件下，器件的电场分布图(a)；能带图(b)

当电压继续增加时，使正电极一边 $x=L$ 处的能带为平带，电场为 0. 整个器件内全部耗尽，而且电场指向同一个方向从右向左，相应的电压称为"平带电压" V_{FB}. 此时电子电流仍很小，但是由于空穴势垒下降，从正偏置处开始有空穴注入. 当电压超过 V_{FB} 时，能带进一步变陡，内部电场增加，直到在反偏置的电极 1 处电场最大点发生雪崩击穿，使电流激增. 图 7.9 绘出了这一情况的电场分布和能带图. 通常器件工作在平带电压 V_{FB} 与击穿电压之间.

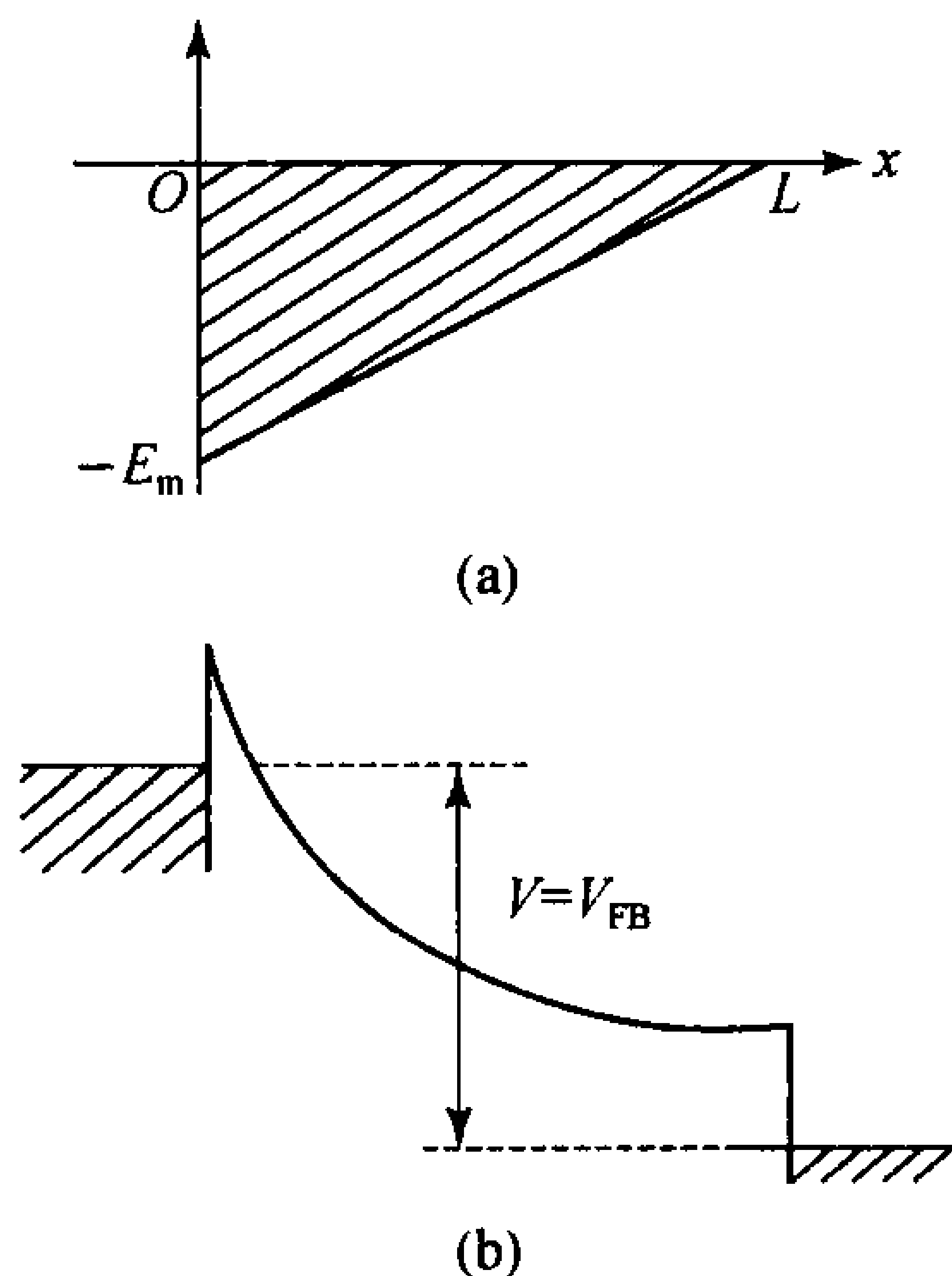


图 7.9　平带电压 V_{FB} 下，器件的电场分布图(a)；能带图(b)

§7.3 硅光电探测器[8]

硅材料能对 850 nm 波段光有很好的响应，因而在 500～1000 nm 波段，体硅光电探测器是理想的选择对象. 目前，研究最多的高速硅基 PD 结构主要有两种：p-i-n 和 MSM 探测器. 对于 p-i-n 结构的探测器，又可以分为垂直结构和水平结构. 由于硅对 850 nm 光的吸收系数很低，只有 0.06 μm^{-1}. 一般的垂直结构硅 p-i-n 探测器的响应度和响应速度将相互制约. 要获得高的响应度，则必须有长的光吸收长度，也就是说，在 p 层和 n 层中间要有厚的低掺杂 i 层，这会使光生载流子的渡越时间增大，器件的响应速度下降. 这一制约关系不解除，则难于制作出高速、具有合适光响应度的硅基 PD. 水平结构的 MSM 探测器和 p-i-n 探测器使光的传播方向与光生载流子的运动方向垂直，从而可以分别控制光吸收长度和光生载流子的渡越长度. 但是一般的水平结构的 p-i-n 和 MSM 探测器的电场分布集中在样品的表面，由表面向内部迅速减小. 虽然它们可以达到高的光响应，但只有在近表面产生的光生载流子可以在强电场作用下迅速到达电极，较深层处的载流子会在低电场的情况下缓慢地到达探测器的电极，对光电流产生贡献，这将使探测器的响应速度大大降低. IBM 的 M. Yang 等[9]采用挖槽的办法较好地解决了这一问题. 他们在硅片上用反应离子刻蚀制作出叉指状的深槽，深度为 7 μm，在表面处槽的宽度为 0.35 μm，指间距离为 3.3 μm. 然后用掺磷(P)和掺硼(B)的非晶硅分别填满深槽，再进行高温退火，使非晶硅结晶成多晶，同时激活掺杂杂质. 并使杂质向硅中作一定的扩散，在离开深槽界面一定距离的硅中形成 p-n 结，最后制作硅化物和金属接触. 深槽的制作在 7 μm 的深度范围内的电场强度均匀，该范围内产生的光生载流子在强电场下能快速漂移到达电极. 这使 PD 的量子效率和响应速度都有所提高，但还是有部分深层的光生载流子要通过扩散的过程才能到达电极，这会影响器件的响应速度. 他们研制的器件在 845 nm 处的响应度为 0.47 A/W，频率响应谱有低频拖尾，6 dB 带宽为 1.5 GHz. 在 670 nm 处，3 V 下的 3 dB 带宽为 2.5 GHz.

为了提高硅光电探测器的速度，必须使光生载流子处于可以使它们达到饱和漂移速度的强电场中，同时采取措施将不在强电场中的光生载流子屏蔽掉，避免这一部分载流子通过慢过程的漂移或扩散到达 PD 的电极，影响其响应速度. 当然，这会牺牲器件的量子效率，所以在量子效率和响应速度之间要作出合理的折中考虑. 一种方法是将探测器制作在硅薄膜上. 图 7.10 为美国科罗拉多大学的 H. C. Lee 等[10]研制的在硅薄膜上的 MSM 结构 PD，硅薄膜的厚度为 3～7 μm，具有纹理结构的不平整下表面是为提高器件的响应度. 更常用的手段是将探测器制作在 SOI 上，利用埋层的 SiO_2，避免在其下面的硅基底内的光生载流

子被探测器收集，影响其速度. 目前报道的最快响应的硅探测器是 Liu 等用 SOI 材料研制的，他们所用的 SOI 材料顶层硅厚度为 100 nm，在 780 nm 下响应度为 5.7 mA/W，带宽为 140 GHz. 虽然速度很快，但低的响应速度使它难于在实际中得到应用. 要提高器件的响应度，可以适当增加顶层硅的厚度，同时牺牲一点带宽.

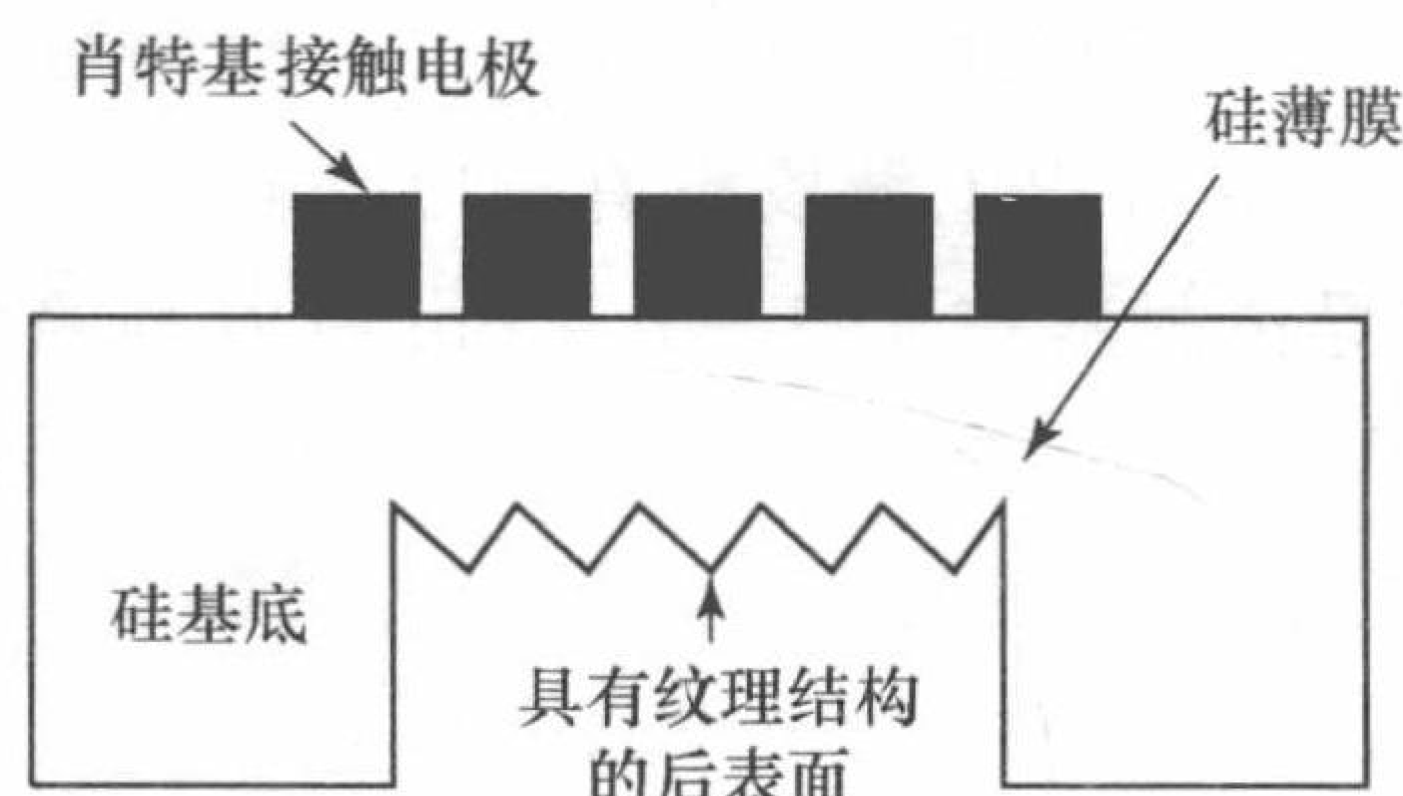


图 7.10 制作在有纹理结构的 Si 薄膜上的高速 MSM 光电探测器的结构示意图

要研制实用化的高速硅基 PD，解除探测器量子效率与响应速度的制约，人们提出了一些新的思路. 其中之一就是制作具有微结构的硅表面，利用光在微结构硅表面处的全内反射，增加光吸收. 如文献[10]研制出了硅薄膜型 MSM 的 PD，他们将薄膜表面用等离子刻蚀成一些微结构，用以增加对光的吸收，提高量子效率. Bell 实验室的 B. F. Levine 等[11]用 SOI 为材料，用高密度等离子源刻蚀表面，使表面粗糙化，形成尺度约 0.1 μm 的微结构，3 μm 厚的硅对进入其中的 880 nm 光的吸收效率超过 50%. 另一个有效方法是制作共振腔增强（RCE）结构，即将光吸收响应介质材料置于 F-P 腔中，符合共振条件的光将在腔中产生共振，被共振增强吸收. 这样，即使是较薄的光吸收材料也能得到较大的量子效率. 最简单的硅 RCE 探测器可以用 SOI 来实现，适当设计顶层硅和 SiO_2 层的厚度，由于硅与 SiO_2 的折射率差较大，在它们的界面处可以形成较大的反射，成为下反射镜，而硅表面与空气界面可以成为另一反射镜，两反射镜之间的硅层为光吸收介质，从而可以实现 RCE PD 结构. 但是 SOI 材料只有一层 SiO_2，形成的底反射镜的反射率低，要充分发挥 RCE 结构的优势，底反射镜的反射率越高越好，所以有必要进一步提高底反射镜的反射率. 美国波士顿大学的 M. K. Emsley 等[12]用键合和智能剥离的方法制作了具有双 SiO_2 层的 SOI 基底，反射率达到 90%以上. 利用该材料研制的 RCE 探测器工作波长为 800 nm，量子效率达到 40%，响应时间为 29 ps，带宽可以超过 10 GHz. 人们还研究了多种制作高反射率底镜的制作方法，以便制作高性能的硅基 RCE-PD，包括合并外延层过生长多次氧注入形成 SiO_2 埋层和硅外延制作多周期 Si/SiO_2 反射镜，乳胶键合与智能剥离相结合，将薄层单晶硅制作在高反射率的 Si/SiO_2 反射镜上.

§7.4 锗硅光电探测器

制作高响应速度、高响应度、低暗电流，响应波长在通信波段 1.3～1.55 μm 的近红外光电探测器，并最终实现高带宽的光电集成接收机芯片一直是人们追求的目标[13～14]. 由于能带结构的固有特性，Si 单晶材料对近红外光存在吸收系数低、吸收长度长、对 1.1 μm 以上波段没有响应等问题.

传统的Ⅲ-Ⅴ族半导体材料在 1.3～1.5 μm 有着高吸收系数，基于Ⅲ-Ⅴ族的光电探测器也已投入商用；但是，Ⅲ-Ⅴ族半导体材料价格昂贵，热导性和热机械特性较差；另外，他们的制造工艺不能和当前的成熟 CMOS 工艺兼容，极大限制了它在集成光电子学中的应用. 锗由于其在单晶状态下的带隙为 0.66 eV，因此可以吸收光波长到 1.87 μm 的光[15]. 另外，Ge 比 Si 的迁移率大，相应电子器件的响应速度大. 同时更为重要的是它的制备工艺和 CMOS 工艺兼容. 硅基 Ge 外延材料将以其优良的加工性、低廉的价格、优良的光电特性、灵活优异的集成性等特点，在微电子学、光子学、光电集成和高效太阳能电池等方面发挥重要作用. 因此研究 Si 基 Ge 探测器引起了人们的极大关注.

7.4.1 Ge/Si 材料的基本物理特性

Ge 与 Si 同属 IV 族元素，尽管硅和锗具有相同的金刚石结构，但它们的晶格常数不同，其中 Si 的晶格常数为 0.5431 nm，Ge 的晶格常数为 0.5657 nm，因此，Si 基底上外延生长 Ge 时，其晶格失配度达 4.2%. 由于 Ge-Ge 键比 Si-Si 键弱，因此 Ge 具有比 Si 小的表面能. 在 Si 上生长 Ge 时，开始时满足浸润条件(即薄膜表面能和界面能之和总是小于基底的表面能)，生长是层状生长，然而随着生长厚度的增加，由于晶格失配，应变能增加，浸润条件不再满足，生长将转化为岛状生长. 所以 Si 基底上生长 Ge 是典型的 SK 生长模式. 因此，一方面，在 Si 上直接沉积纯 Ge 层所形成的 Ge/Si 异质结构材料将会不可避免地遇到很大应力，积聚应变能. 其能量随着厚度增加而线性增长. 当应变层达到一定的临界厚度时，积聚在应变层中的应变能就将会以位错或者表面起伏的形式释放出来，这称为应变弛豫[16,17]. 位错为电子空穴对提供复合中心，故而提高电子空穴对的复合几率，于是暗电流变大，最终减低光电探测器的性能；另一方面，为了获得较高的响应度，我们又不得不需要一定厚度的 Ge 吸收层. 因此，必须尽量拓展 Ge 吸收层的临界厚度，这样所得到的较厚的高质量 Ge 层，便可解决暗电流与响应度之间的矛盾问题. 根据 R. People[18] 等的动态模型，得到 Ge_xSi_{1-x}/Si 材料的临界厚度的理论计算公式

$$h_c = \frac{2.34 \times 10^{-3}}{f_m^2} \ln\left(\frac{h_c}{0.4}\right), \tag{7.4.1}$$

其中 f_m 是失配率，临界厚度 h_c 的单位是 nm. 同样地，在 Ge_xSi_{1-x}基底上沉积一层 $Ge_ySi_{1-y}(y>x)$，临界厚度的计算同样也如上式(7.4.1)，不过 f_m 改成了相应基底与吸收层的失配率. 理论计算表明使用 Ge_xSi_{1-x}代替 Si 作为 Ge 吸收层的缓冲层的话，这将会显著地改善提高 Ge 的质量和厚度，对探测器的性能有着重要提升作用.

麻省理工学院的 Ishikawa 等[16,19]发现到了直接生长在 Si 上的 Ge 带隙缩小，光吸收增强的现象，他们认为这是由 Ge 和 Si 的热膨胀系数不同产生的. Ge 的热膨胀系数比 Si 大，当高温生长的 Ge 层冷却到室温时，Ge 就会受到张应力，而应力会影响能带结构，改变 Ge 的带隙. 一般当张应力 $\varepsilon_{\parallel-G_e}=0.2\%$时，直接带隙会从 0.8 eV 缩小至 0.77 eV，Ge 材料的带隙变小将增强对光的吸收，增大 Ge 的吸收系数，扩展吸收截止波长. 这对于提高光电探测器在长波长段尤其是 1.55 μm 波段的响应度有很大裨益.

7.4.2　Ge 材料制备方法

Si 基底的高质量 Ge 层生长一直是高性能近红外光电探测器制造的基础. 不过由于晶格失配和热膨胀失配导致的高粗糙度和高位错密度一直是我们所面临的一大挑战，限制了探测器性能的提高. 在 1984 年，Luryi[17]等利用分子束外延技术生长 Ge 材料，开始了硅基 Ge 探测器的研究. 他们利用渐变 Ge_xSi_{1-x}材料作为缓冲层制造了 n-i-p 结构的 Ge 探测器，位错密度达到 $10^9\ cm^{-2}$，暗电流密度 50 mA/cm^2. 但从那时起，进一步提高 Ge 材料质量的生长遇到了巨大困难，Ge 探测器的性能提升一直停滞不前. 直至最近生长技术和设备取得了一大步进展，Ge 探测器又迅速发展起来. 现在有三种主要方法制造 Ge 材料吸收层：外延技术、低温晶片键合及离子切割技术和快速融化生长技术.

1. 外延技术

目前在 Si 基底上外延生长 Ge 材料的主要工艺有三种[20]：① 组分渐变的 SiGe 缓冲层(Buffer)工艺. 该工艺首先生长 Ge 组分从 0 到 100%逐渐增加的 SiGe 缓冲层，使应变逐渐释放，以获得位错密度低的缓冲层，然后在其上生长 Ge 外延层. 使用该方法可以生长出晶格质量很好的 Ge 材料，位错密度可以达到 $10^6\ cm^{-2}$数量级，但是由于表面会有很大的起伏，必须在生长后或生长中间插入化学机械抛光工艺流程，制作的工艺复杂耗时，而且为了获得好的晶体质量，以 SiGe 组分渐变层的厚度将达到 10 μm 以上，这样的材料不利于制作集成器件. ② Si 图形基底上生长 Ge. 就是在刻蚀有图形的 Si 基底上进行 Ge 的生长，主要有两种方式：一种是在 Si 基底上刻蚀出一维或二维结构的台面，然后进行

Ge 的外延生长;另一种图形基底是在 Si 基底上制备 SiO_2 薄膜,然后光刻并刻蚀 SiO_2 露出生长 Ge 的窗口,Ge 将选择性地在露出 Si 的位置生长,并可以横向过生长而在 SiO_2 表面合并,形成完整的 Ge 外延层. ③ 低温 Ge 缓冲层工艺. 该工艺首先在 400 ℃以下的温度下生长出应力弛豫的 Ge 缓冲层,厚度约 50 nm,然后将基底温度提高到 600 ℃左右,生长合适厚度的 Ge 层. 生长后,为了提高材料质量,可以进行循环退火处理. 最终获得的材料的位错密度一般在 10^7 cm^{-2} 数量级的水平,表面的平整度也比较好[21]. 该方法的优点是工艺简单、生长时间短、缓冲层薄、适合制作集成器件. 根据这一原理,人们提出了化学气相沉积(chemical vapour deposition, CVD)两步生长 Ge 的方法,即低温 Ge 缓冲层方法. 由于 CVD 方法生长 Ge 时,在低温时表面会有 H 的覆盖,第一步的低温过程中 Ge 的生长将保持二维生长,并且以位错而不是以起伏的形式释放应力,从而获得平整弛豫的 Ge 缓冲层. 接着在缓冲层上约 600 ℃下生长厚的 Ge 材料. 目前人们基本上倾向于用 Ge 低温过渡层技术来外延生长硅基 Ge 材料,并取得了很好的结果. 典型外延技术生长的 Ge 材料及其性能对比如表 7.1 所示.

表 7.1 典型外延技术生长的锗材料及其性能对比

研究小组	设备	缓冲层技术	缓冲层厚度	Ge 层的厚度	表面粗糙度/nm	位错密度/cm^{-2}
Thomas 等[22]	LEPECVD	组分渐变的 GeSi 缓冲层	10 μm	1 μm	3.3	1.1×10^5
Luo 等[23]	UHVCVD	双 GeSi 层	2×0.8 μm	1 μm	3.2	3×10^6
Luan 等[24]	UHVCVD	低温生长 Ge	30 nm	1 μm	1~2	2.3×10^7
Loh 等[25]	UHVCVD	低温生长 GeSi +低温生长 Ge	30+30 nm	100 nm	1.4	1.6×10^6
Vanamu 等[26]	CVD	2D 图形基底	—	10 μm	<2	5×10^5
Vivien 等[27]	RPCVD	低温生长 Ge	50 nm	280 nm	—	3×10^6

应用低温缓冲层技术生长 Ge 材料时,有两个关键点需要注意: 一是温度控制必须仔细;二是缓冲层的厚度必须认真把握. 这两个因素对于最后形成的纯 Ge 材料质量有着重要决定作用. 低温缓冲层技术目前使用较为普遍,但是这个方法所带来的位错无可避免,进一步降低位错的空间不是很大,同时,由于沉积高温 Ge 材料吸收层的温度可能高于 450℃,因而与 CMOS 后端工艺不兼容.

2. 低温晶片键合及离子切割技术

在 2008 年,Chen[28,29] 等利用低温晶片键合及离子切割技术在亚微米硅波导上成功地沉积出高质量的 Ge 材料,并由此制造了 Ge 探测器. 由于获得完全没有因晶格失配而带来的位错缺陷的高质量 Ge 薄膜,因此所得 PD 的暗电流性能突出;不但如此,整个制造过程的操作温度在 400 ℃,这完全与 CMOS 微电子的后端过程[30]相兼容,因此相较于低温高温分子束外延有一定的优势. 具体流

程如下：首先，在 400 ℃条件下用等离子体增强化学气相沉积(plasma enhanced CVD, PECVD)的方法在 4 英寸的 Ge 片上覆盖一层 100 nm 的 SiO_2，随之注入 H^+ 离子用以切割；SiO_2 覆盖层是用来保护 Ge 表面不受离子注入损害的，在注入后即用稀释的氢氟酸去掉. 其次，准备另一片 4 英寸的 SOI 硅片，在经过一系列的步骤(主要是形成特定结构如波导等)之后，将这两片清洗干净，在室温条件下用向下的压力键合起来，然后在氩(Ar)气氛围中退火，温度控制在 100℃，一定的气压增强键合力度，最后温度逐步上升到 400℃. 在如此逐渐上升温度条件下，注入的 H^+ 离子在注入深度处形成微腔，诱导产生的压力将注入深度以上的 Ge 薄膜与下面一整块晶片分离开来[28,29]. 接下来便是具体 Ge 层结构的塑造，如下图 7.11 所示.

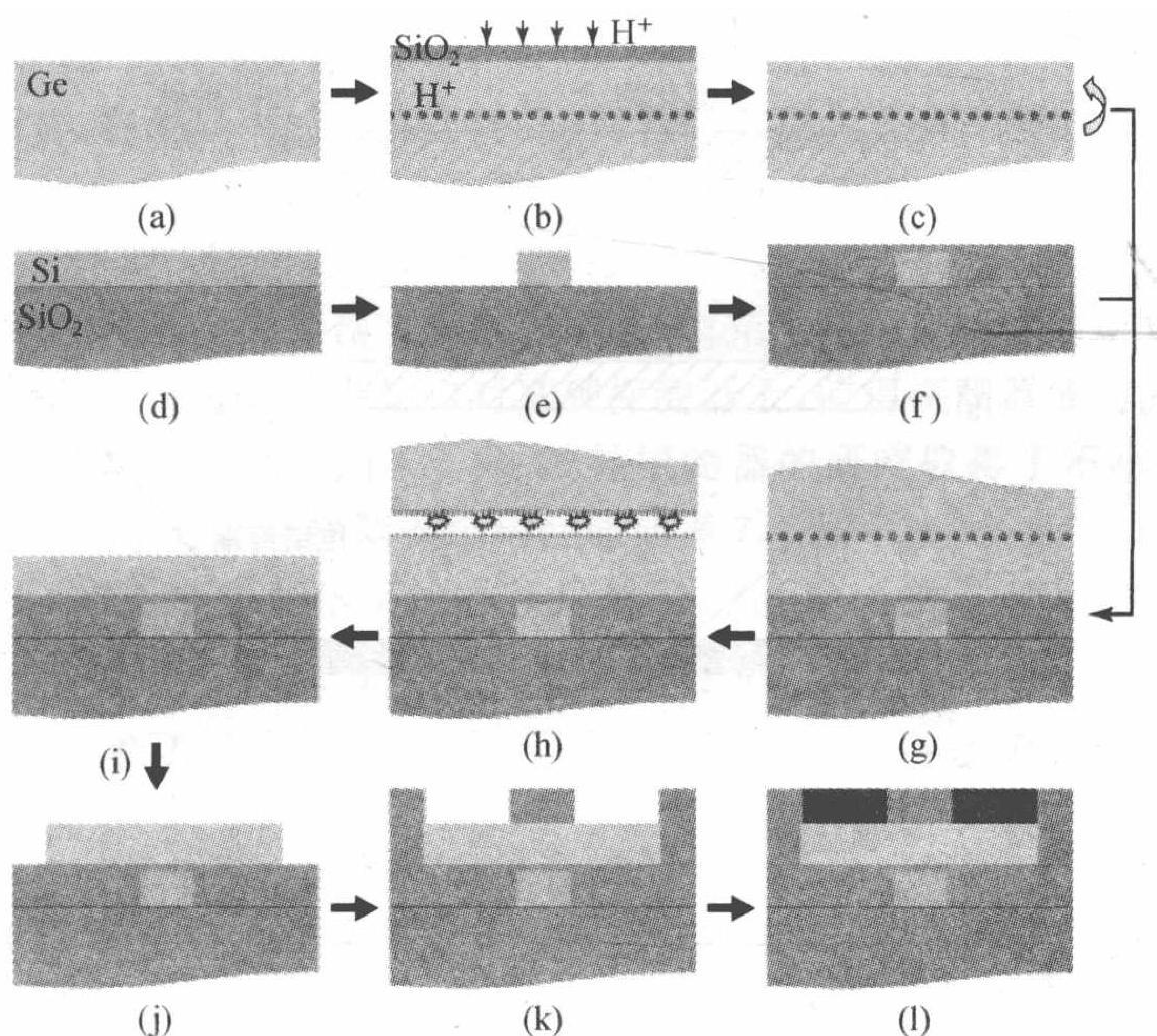


图 7.11　低温晶片键合及离子切割技术流程

通过运用这种方法所生长的 Ge 材料没有位错的出现，由此制造出的相应器件所得到的暗电流测试结果低至 100 nA，十分突出；不过这种方法还是稍复杂，规模生产困难.

3. 快速融化生长技术

快速融化生长技术或者液相外延技术[31]是另一个实现无缺陷的薄 Ge 单晶材料生长的方法. 这种方法利用了液相外延技术原理和缺陷瓶颈技术，而且这个高质量的单晶结构能够与硅基的制造相兼容. 因此这项技术可以很方便地在任意位置集成绝缘体上锗(Ge on insulator, GOI)薄膜，形成一些 Ge 器件如 Ge 探

测器等.具体流程如下：首先，在硅基底上沉积一层硅氮化物作为隔离层(氧化物是另一个选择)；然后，利用光刻和刻蚀技术在氮化薄膜上形成一些特定图案作为种子窗口.其次，将 Ge 材料无选择性的溅射在基底上，既覆盖氮化物层又覆盖暴露出来的硅层.再次，Ge 薄膜被塑造成期望的图案，然后在 300 ℃下用低压化学气相沉积(low-pressure CVD, LPCVD)沉积低温氧化物层(low temperature oxide, LTO).用快速热退火方法将硅片加热到 940 ℃保持 2 s，其中温度上升率为 100 ℃/s.在退火过程中，Ge 材料会融化(因为超过了 Ge 材料的熔点)，LTO 层充当了微型坩埚作用，盛起了 Ge 液体使之不会到处流动.不过硅基底却在这样的温度下仍然保持着固相.自然冷却下来后，便会发生液相外延：Ge 就会从 Si/Ge 交界面的种子窗口(seeding window)一直生长到氮化膜的顶层[31]，如图 7.12(a)，(b)所示.

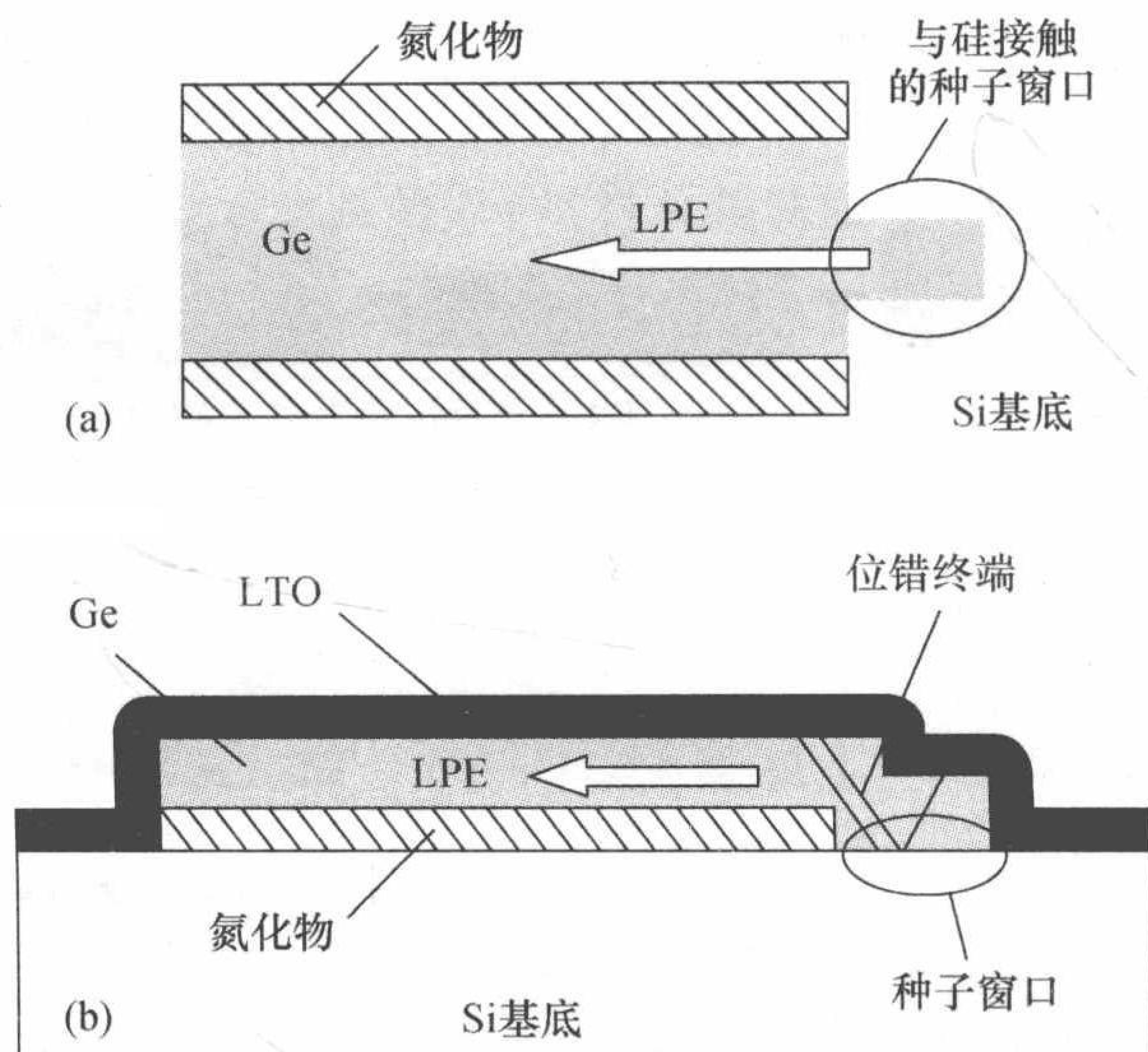


图 7.12　快速融化生长示意图

这项技术仅仅需要一些快速退火系统和整片晶片加热技术，因此相对于低温分子束处延(molecular beam epitaxy, MBE)和低温晶片键合及离子切割技术而言，它更简单，更快捷；不过它也有一些缺点：如高温融化可能与 CMOS 后端工艺不容.

传统的光电探测器(Ⅲ-Ⅴ族探测器)的结构属于典型的吸收层、电荷层和倍增层分立(separate absorption, charge and multiplication, SACM)多层结构，但这是不适合以 Ge 材料为主的新型光电探测器.因为 Ge 材料与 Si 晶格失配，生长多层而且质量高的 Ge 层则是不太现实.因此为了提高 Ge 探测器的性能，新颖结构的提出迫在眉睫.

7.4.3 锗硅波导光电探测器

波导型探测器在保持吸收区厚度的前提下可以大大增大吸收长度，将探测器设计成波导型，使用锗探测器作为波导的一部分，由于光的传播和吸收沿着波导方向，而载流子输运则沿着与之相垂直的方向，这种结构允许波导集成的光电探测器显示出高的速率同时可以达到大约 100% 的内量子效率. 另外，波导集成的光电探测器的器件面积可以比自由区域的光电探测器小 10 倍，所以对于同样的暗电流密度来说，它的绝对暗电流是非常低的. 由于噪声是由绝对暗电流（而不是暗电流密度）决定的，所以这种结构的探测器的噪声是很低的. 同时波导集成也改善了锗光电探测器的灵敏度. 由于波导集成型锗硅探测器具有以上的优点，因此近年来对这种类型的探测器的研究引起了人们的极大关注.

1. p-i-n Ge 波导探测器

一般而言，将波导中的光耦合到 Ge 材料吸收层主要有两种方法：倏逝耦合(evanescent coupling)和对接耦合(butt coupling). 因此，波导探测器可能有三种结构：第一种是把波导集成到锗探测器的顶端，光以倏逝波的形式耦合到锗探测器上；第二种是在探测器下面制作出波导，光也是以倏逝波的形式耦合到探测器上；第三种结构是把波导端口对准到探测器上，使得探测器成为波导向外延伸的一部分. 近年来，人们对 p-i-n Ge 波导探测器的研究取得了不小的进展，目前报道的不同的 p-i-n Ge 探测器结构性能如表 7.2 所示.

2007 年，Ahn 等[32]制作了一种倏逝耦合的波导探测器，其结构如图 7.13 所示. 它的 Ge 层厚度是 1 μm，使用了倏逝耦合方式，在 −1 V 下有着 1 μA 的暗电流，在波长是 1.55 μm 处有着 1.08 A/W 的响应度，这些都是相当优异的；不过其致命的弱点就是带宽比较小，只有 7.2 GHz，这并不能满足人们目前的需求. 通过分析，一个可能原因是它并没有使用 SOI 基底，一些电子空穴对可能会扩散出电场驱动区域，在被金属电极收集前随机分布，因此降低了带宽. 另一个可能原因是较大的电容带来了较大的 RC 延迟，最终同样降低了带宽. 此后，很快诞生了诸多改进方案.

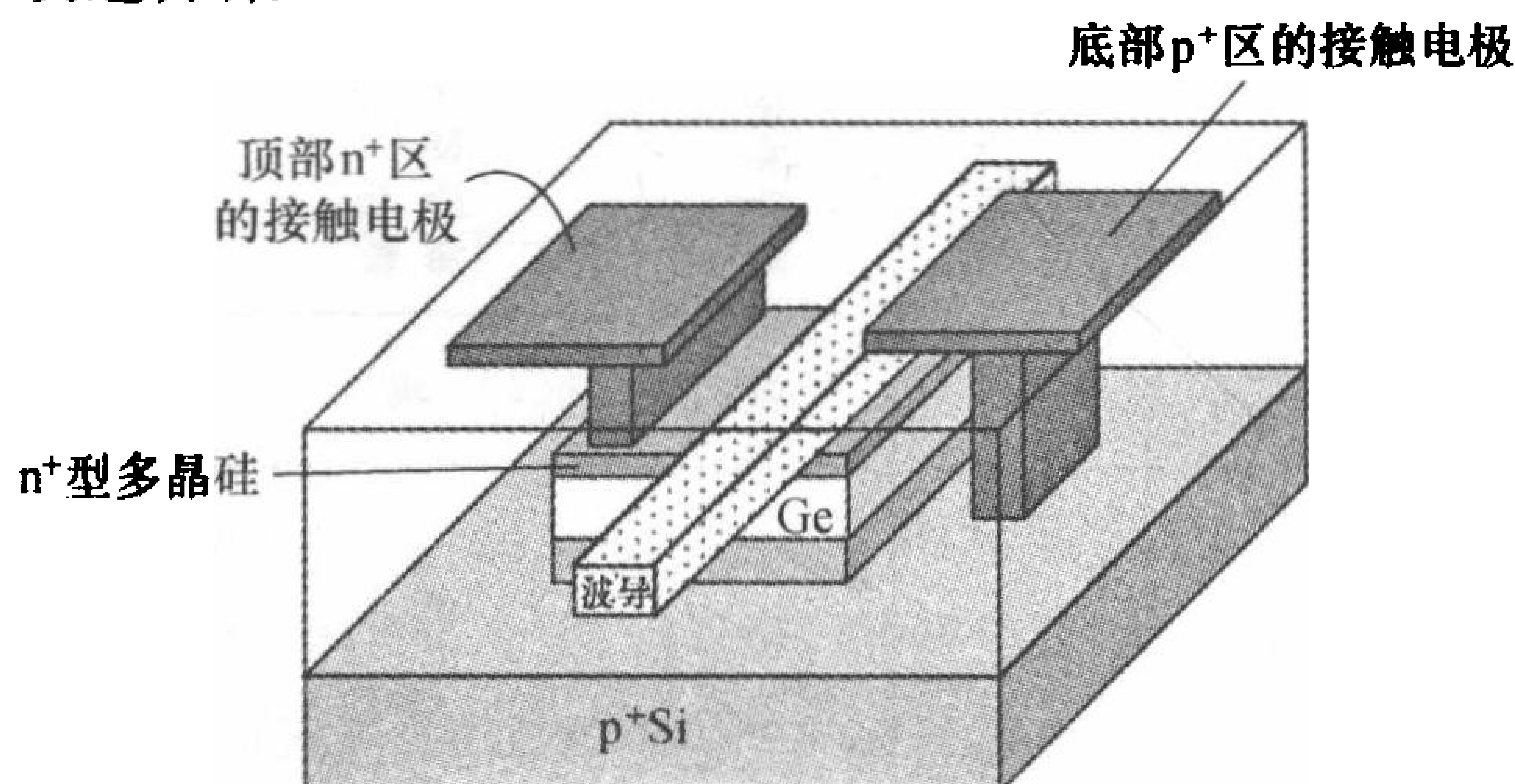


图 7.13　波导集成的 Ge p-i-n 探测器的结构示意图

表 7.2 不同 p-i-n Ge 探测器结构性能比较*

研究小组	SOI 基底	耦合结构	尺寸 d(厚度)/μm, A(对于表面积而言,宽度×长度)/μm²	暗电流 I_{dark}/μA 暗电流密度 I_{dark}/(mA/cm²)	响应度/(A/W) 外量子效率	电容/fF	−3 dB 带宽 /GHz
Samavedam 等[15]	不是	波导垂直的倏逝耦合	d=1.106 μm A=7×10 μm²	1@−1 V; 0.06@−0.1 V	@−1 V,λ=1.55 μm 1.08;90%	≈400	7.2@−0.1 V~−1 V
Tersoff 等[16]	是	Rib 波导垂直的倏逝耦合	d=0.7 μm A_1=4.4×100 μm² A_2=7.4×50 μm²	@−2 V, A_1:267 nA;74 A_2:169 nA;51	@−2 V,λ=1.55 μm A_1:1.16;93% A_2:0.89;71%	A_1:72.8 A_2:66.7	@−2 V, A_1:29.4 A_2:31.3
Chen 等[29]	是	波导倏逝耦合	d=0.11 μm A=4.4×80 μm²	@−1 V, VPD:0.44 LPD:0.74	@−1 V, VPD:0.29 LPD:0.81	—	@−1 V, VPD:4.4 LPD:2.8
Fedeli 等[30]	是	波导横向的倏逝耦合	d=0.22 μm w=2.4 μm 电极间距=0.8 μm	15 nA@l=5 μm 30 nA@l=10 μm 45 nA@l=15 μm 60 nA@l=20 μm	@−1 V, λ=1.55 μm; 内量子效率 0.4@l=5 μm 0.5@l=10 μm 0.58@l=15 μm 0.65@l=20 μm	0.8(耗尽区)	18@−1 V
Tomp 等[17]	是	波导垂直的对接耦合	d=0.34 μm A=3×15 μm²	0.018@−1 V 1.3@−2 V 33@−4 V	λ=1.55 μm 1@−1 V~−4 V	—	12@0 V 28@−2 V 42@−4 V
Feng 等[37]	是	Rib 波导横向的对接耦合	d=0.8 μm A=2.4×10 μm²	0.24@−0.5 V 1.3@−1 V	λ=1.55 μm 1.06@TE 1.1@TM	—	17.5@0 V 32.6@−1 V 36.8@−3 V

注：@表示当满足@后面的条件时，将得到@前面的结果.

同年，Intel公司[33]相较于前者作出了如下几项改动：① SOI基底的使用增强了响应速度，大大降低了暗电流；② 脊波导的运用带来两个好处：一是脊波导的光场限制与矩形波导类似，二是脊波导更容易引出金属收集板；最后，为了增加耦合效率，他们制造了一个楔形耦合器来增强了对光场的吸收. 尽管Ge层的厚度减小，在同样偏置下电场强度更大，但是暗电流还是小于1 μA，响应度则保持在1 A/W；不过最明显的变化则是因较薄的Ge层和很小的电容所取得的相对很高的带宽(31 GHz)，其结构如图7.14所示.

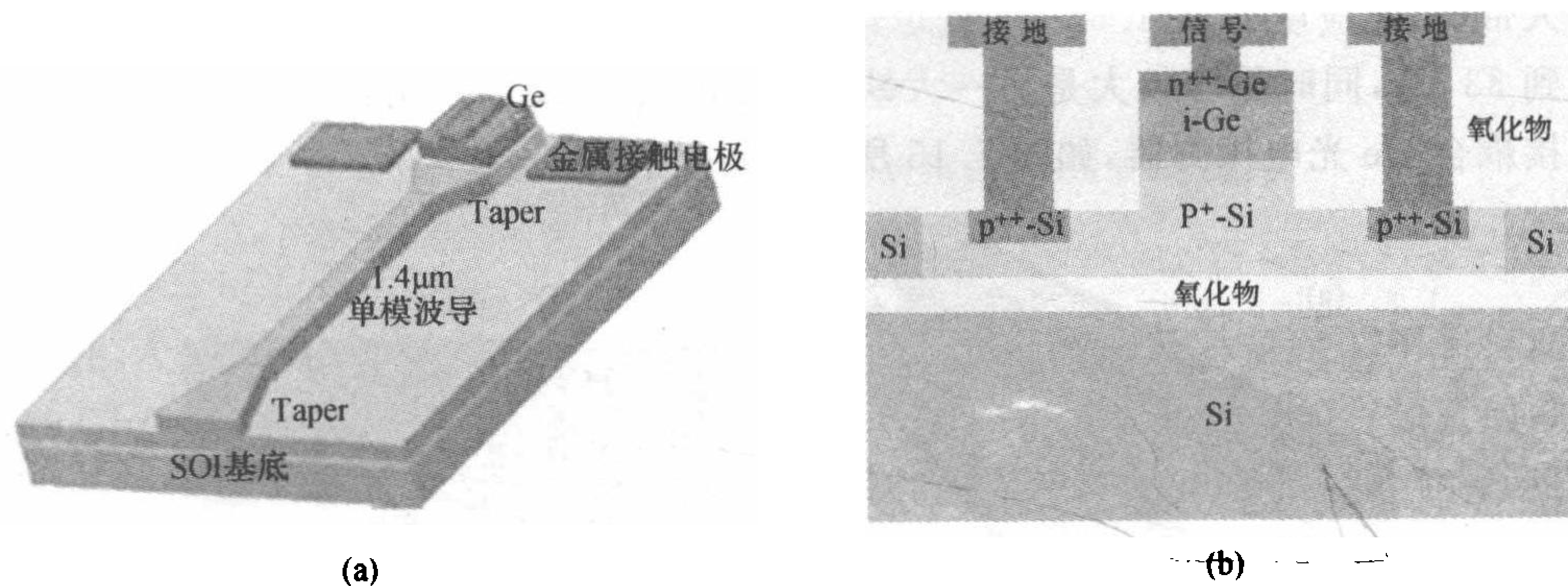


图7.14　Intel公司Ge探测器结构示意图

该图也是典型的p-i-n探测器结构，其带宽的理论计算如下：因渡越时间产生的带宽为

$$f_{\text{transit}} = \frac{0.45\, v_{\text{sst}}}{d}, \tag{7.4.2}$$

因RC延迟产生的带宽为

$$f_{\tau} = \frac{1}{2\pi RC}, \tag{7.4.3}$$

电容由下式决定：

$$\left(\frac{A}{C}\right)^2 = \frac{2}{qN\varepsilon}(V + V_{\text{bi}}) + \frac{d^2}{\varepsilon^2}, \tag{7.4.4}$$

因此总的3 dB带宽为

$$f_{3\,\text{dB}} = \frac{1}{\sqrt{f_{\text{transit}}^{-2} + f_{\tau}^{-2}}}. \tag{7.4.5}$$

在该结构中，700 nm的Ge层厚度使得相对较低的偏压($V_{\text{r}} = -2$ V)仍能产生很高的电场，由此算得的带宽值与实际测量值间相差不大. 不仅如此，若以此公式对带宽进行优化设计，则可得到关于器件结构的d和A等参数值，有着一定的指导价值.

2008 年，新加坡的 Wang[34,35] 等针对波导侧向和垂直倏逝耦合两种方式设计了一系列的结构比较两者的优劣. 基于研究结果，他们认为侧向的耦合方式要优于垂直方式. 不过响应度不够高、带宽不宽始终是他们的缺点.

除了倏逝耦合，2009 年的两篇报道使得对接耦合成为一大焦点. 由于纯 Si 和纯 Ge 的折射率差别不大，因此光反射影响较小，光场分布相似，这便为对接耦合提供了前提条件. 相较于倏逝耦合方式，对接耦合更简单，效率更高. Vivien[36] 等设计了脊波导垂直对接耦合 Ge 光电探测器，在 −4 V 的偏压下达到了 42 GHz 的惊人带宽，响应度 R 在 1.55 μm 处也达到 1 A/W，然而暗电流却也成指数上升达到 33 μA，同时偏压较大是另一大缺点. Feng 等[37] 设计和制备了脊波导侧向对接耦合 Ge 光电探测器，如图 7.15 所示.

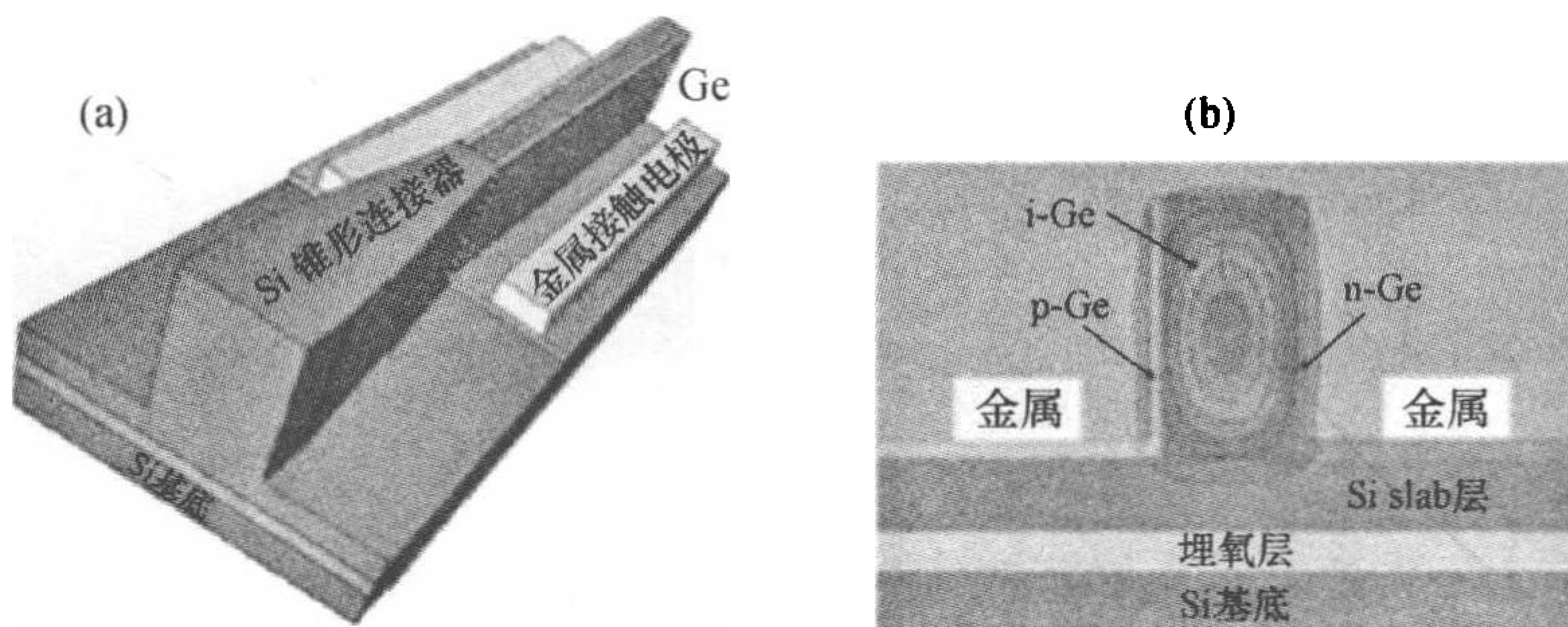


图 7.15 大截面 SOI Ge 探测器结构示意图

这种结构的性能非常好：暗电流较低(在 −1 V 偏压下为 1.3 μA)，响应度对偏振不敏感，在 1.55 μm 处，TE，TM 下均达到 1 A/W 以上，同时在 −1 V 偏压下的带宽超过 32 GHz，这可以使器件在更高速率的条件下工作. 这些好的性能对于 Ge 探测器的应用又向前推进了一步. 不过在垂直方向上沉积 2.4 μm 厚的 Ge 材料将会耗时过长. 另外，对 Ge 层的侧向掺杂也有一定的难度.

为了进一步改进波导探测器的响应度和带宽，2011 年，Kotura 公司提出了两种方法来实现[38]，其结构图如图 7.16 所示.

两种结构的 Ge 层的长度均为 10 μm，结构 A 的 Ge 层的宽度是 1.6 μm，且 Ge 的顶端只有一个金属电极；结构 B 的 Ge 层的宽度是 6 μm，且 Ge 的顶端有两个金属电极. 双重电极可以通过增大电极接触面积来减小接触电阻. 为了减少金属电极对光场的吸收，他们适当增加了 Ge 材料的厚度来减小光场模分布和金属接触部位的重叠. 通过这些设计，在 −1 V 偏压下，A 结构和 B 结构的暗电流分别为 2.6nA 和 11 nA，对应的暗电流密度分别为 49 mA/cm^2 和 22 mA/cm^2. 同样在 −1 V 偏压下，A 结构在波长是 1520 ～1570 nm 范围内的响应度达 0.95 A/W. B 结构在波长 1550 nm 处的响应度是 1.05 A/W，在 1530 nm 处可达 1.19 A/W.

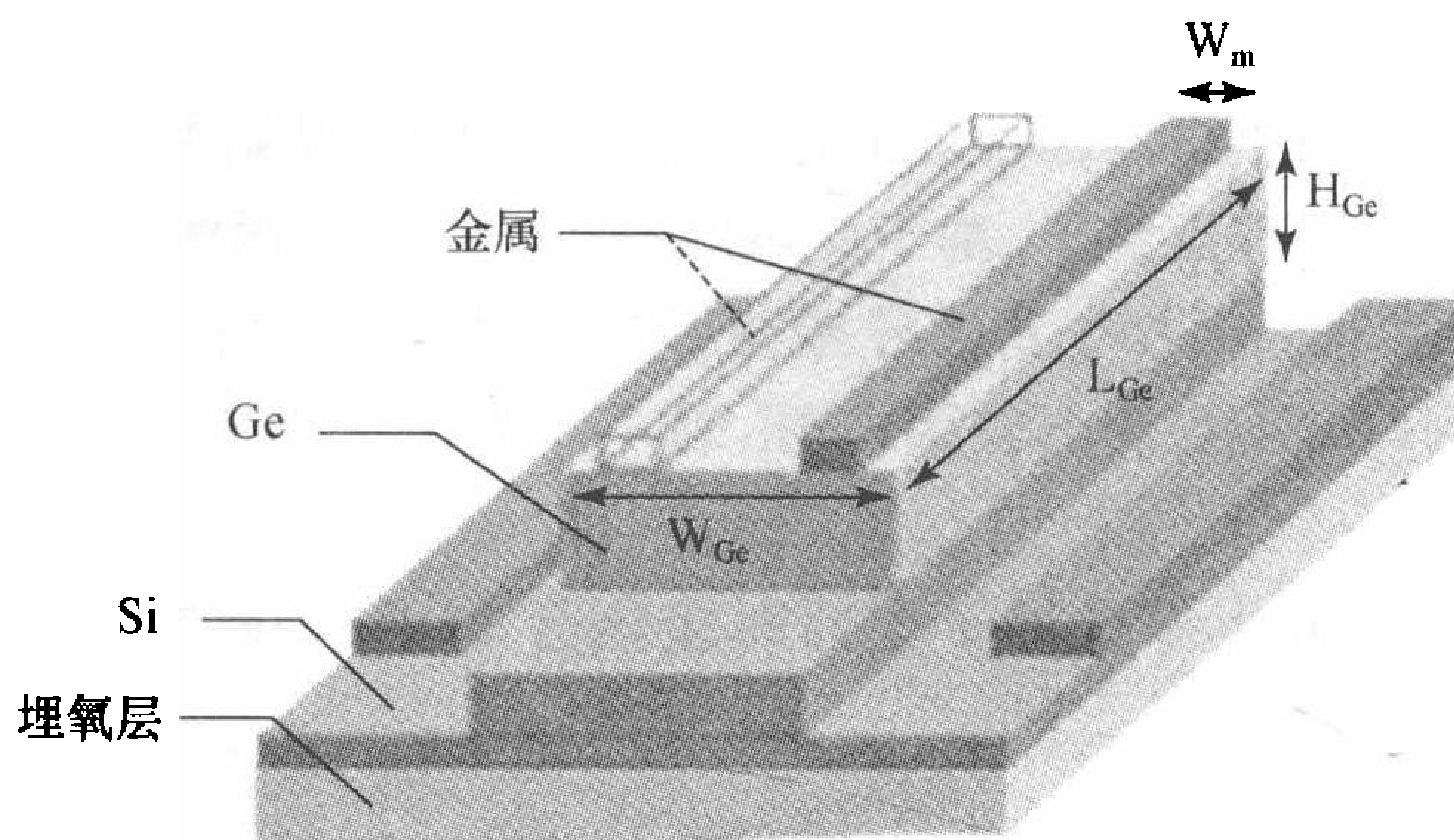


图 7.16 集成在 SOI 波导上的垂直结构的 Ge p-i-n 探测器的三维结构图

另外,A 结构在−1 V 偏压下,所测得的 3 dB 带宽达 36 GHz,这足够用来检测 40 Gb/s 的高速光信号.通过他们的仿真分析,可以通过优化器件的有源区的面积和金属的宽度,来把带宽提高到 40 GHz,同时不会引起响应度的明显变化.这种器件可以很容易地和互阻放大器(transim pedance amplifer, TIA)集成在一起来制成高速、高性能的接收机.

总的来说,对于人们对 p-i-n Ge 探测器的研究已经取得了不错的成绩,但仍存在着很大的空间进一步提高器件的性能,更好的满足人们的需求.

2. MSM Ge 波导探测器

在 p-i-n 结构中,通过掺杂在 Ge/Si 中形成 p/n 区,然后在重掺杂区域,金属与之接触作为电子空穴对的收集板,这种方式为欧姆接触,电阻小为一大特点,所形成的便为 p-i-n 结构;若是金属和未经掺杂或掺杂浓度很低的半导体区域直接相接触,这样在金属和半导体交界区域便产生肖特基势,形成 MSM 结构.当入射光子被材料吸收,产生的电子-空穴对在场强作用下向着极板区域移动,克服肖特基势被金属吸收产生光电流,完成 MSM 光电转换,形成了 MSM 光电探测器.

早期的 MSM Ge 光电探测器如 p-i-n 结构般同为表面入射型,这除了有相同的入射光被反射的困扰外,较多的电极阴影效应更使人头痛不已,由于电子空穴对需要克服较大的肖特基势,因此在较低偏置条件下,正负电极间距很小(200～300 nm)以形成强电场,众多电极形成交叉的叉指,这使得金属电极反射了很大部分入射光,限制了响应度的提高,如文献[39]所述,其量子效率的理论值也就因而不高,如下式:

$$\eta = (1-r)\left(\frac{L}{L+W}\right)(1-e^{-\alpha d}), \qquad (7.4.6)$$

其中 W 为每个叉指的宽度，L 为指间距，α 为吸收系数，d 为 Ge 层厚度，r 表征反射系数. 这很不利于探测器性能的进一步优化设计. 但如果我们采用波导耦合方式，这些问题就迎刃而解了. 基于此，最新的 MSM Ge 探测器大部分都采用波导耦合方式来实现. MSM Ge 波导探测器的制造由于没有较为繁琐的掺杂过程，故而相对 p-i-n Ge 探测器较为简单，表 7.3 列出了不同 MSM Ge 探测器结构性能比较.

2007 年，Vivien[27] 等设计了侧向对接耦合 MSM Ge 探测器，其结构如图 7.17所示.

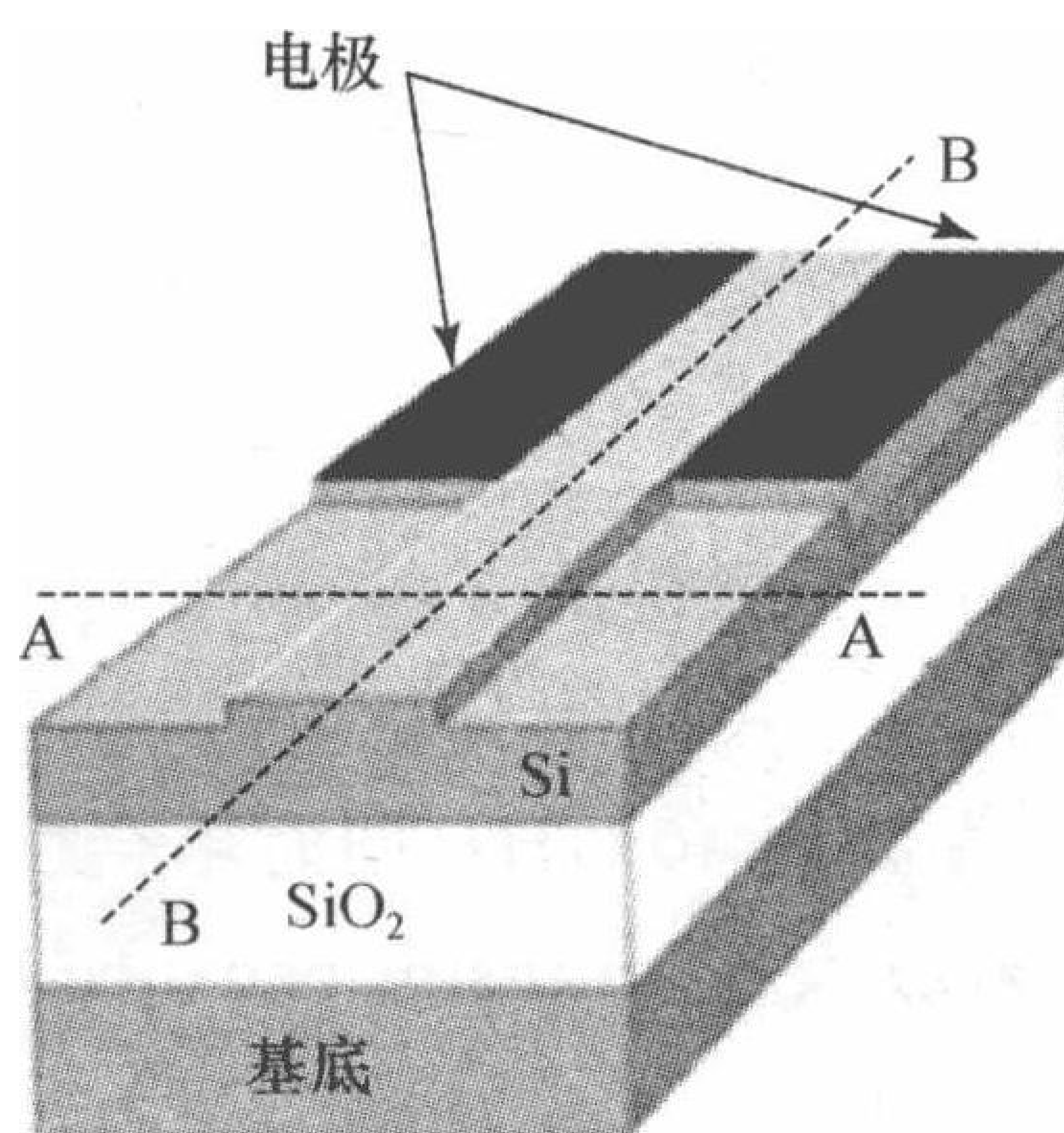


图 7.17 SOI 脊形波导侧向对接耦合结构示意图

采用脊形波导结构同样是便于引出电极，同时当 Ge 外延生长所需的薄 Si 层厚度不超过 50 nm 时，SOI 波导和 Ge 输入面的导模重叠度高达 90%以上. 这样的设计虽然使得响应度很高(1 A/W)，但暗电流过大(在 −1 V 偏压下为 130 μA)，带宽不够(在 −1 V 偏压下为 10 GHz)成为重要弊端.

2008—2009 年，Chen[28,29] 等设计了侧向倏逝耦合 SOI 波导 Ge 探测器，如图 7.18 所示. 在制造过程中，采用了低温晶片键合及离子切割技术沉积高质量 Ge 材料，有效地降低了暗电流. 不过，与对接耦合相比，倏逝耦合的响应度比较低.

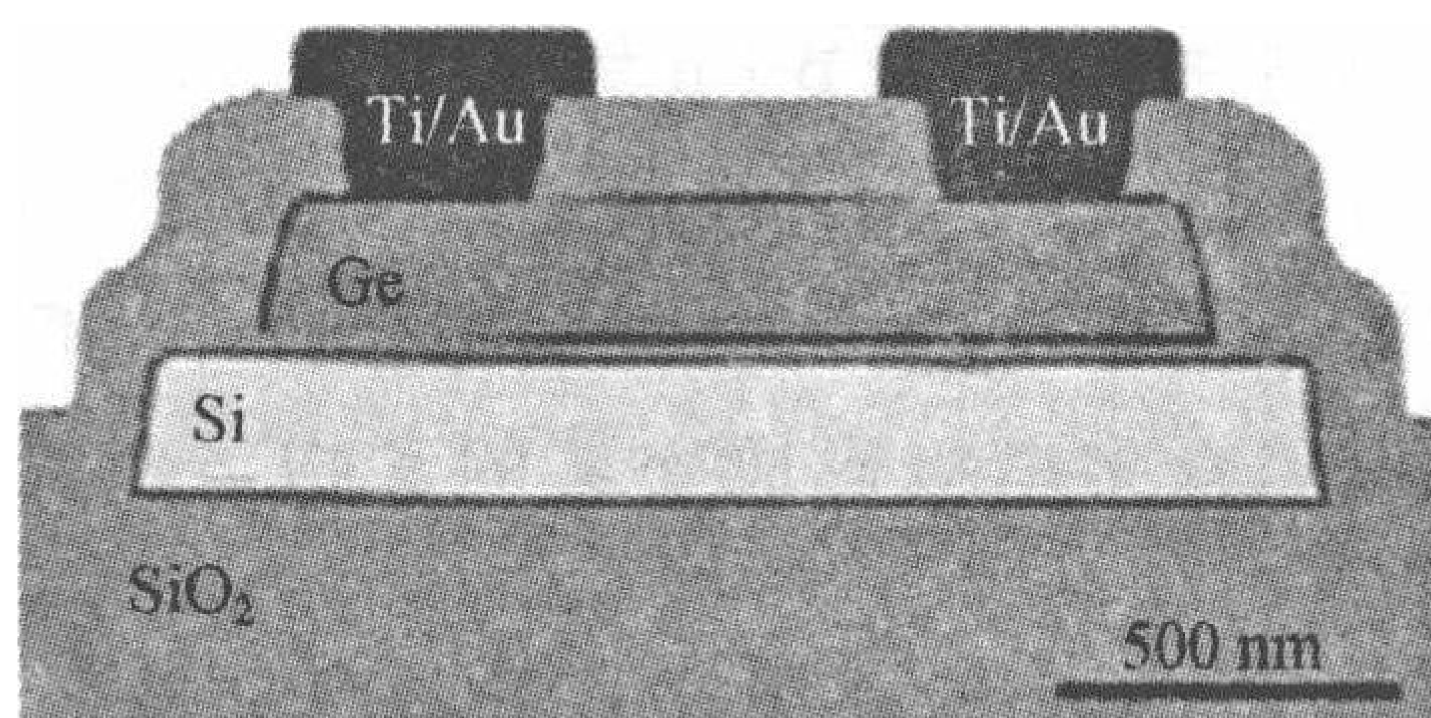


图 7.18 SOI 波导侧向倏逝耦合结构示意图

表 7.3　不同 MSM Ge 探测器结构性能比较

研究小组	SOI 基底	结构(耦合)	尺寸($D,A=W\times L$) 电极间距(S)	暗电流 $I_{dark}/\mu A$	响应度/(A/W) 外量子效率(EQE) 内量子效率(IQE)	电容/fF	−3 dB 带宽/GHz
Luryi 等[14]	SOI 基底 Rib 波导	横向对接耦合	D=330 nm S=1 μm $A=1\times 10\ \mu m^2$	130@−1 V 300@−5 V	λ=1.55 μm 1±0.2 80%	RC<4ps	8.5@−0.5 V 10@−1 V 15@−2 V 17@−3 V 25@−6 V
Liu 等[31]	SOI 基底波导	横向倏逝耦合	D=330 nm S=2.5 μm $A=5\times 80\ \mu m^2$	0.1@−1 V	λ=1.55 μm 0.76@−1 V 60%	—	12@−1 V
Ahn 等[32]	SOI 基底波导	横向倏逝耦合	D=250 nm S=750 nm L=30 μm	0.1@−1 V～−4 V	0.44@λ=1.527 μm 0.3@λ=1.55 μm	2	45@S=500 nm 25@S=1 μm
Yin 等[33]	SOI 基底波导	横向倏逝耦合	D=260 nm S=600 nm $A=1.5\times 30\ \mu m^2$	4@−5 V	—	2.4	50@−5 V
Vivien 等[36]	SOI 基底波导	倏逝耦合	D=100 nm S=200 nm W=500 nm	90@−1 V	V_r=−1 V 0.42@λ=1.31 μm EQE=39% 0.14@λ=1.5 μm IQE=12±3%	10±2	35@−2 V S=300 nm

总之,MSM Ge 探测器面临着两大问题,一是暗电流过大.这一方面可以通过改进沉积 Ge 材料的方法来消除,但可能存在与 CMOS 工艺兼容问题,另一方面通过在金属-Ge 交界处加入一层肖特基势垒增强层来遏制暗电流[40],不过会增加制造复杂度;二是倏逝耦合响应度 R 太低.这两个问题始终影响着 MSM Ge 探测器的进一步的突破.

3. 雪崩型 Ge 波导探测器

传统的雪崩型探测器(主要是Ⅲ-Ⅴ族 APD)需要很高的偏压(−20 V)以实现雪崩放大效应.这虽然可以取得较高的增益,但是会不可避免地遇到噪声放大的不利后果.

由于 Ge 探测器不太可能制造成 SACM 多层结构,同时一般要求偏压较低(−5～−1 V),因此在实验中我们很难观测到雪崩增益.不过非一致性场 APD 是一个重大突破.通过产生非一致性的电场,碰撞电离区域缩短至 30 nm,并能有效减少噪声.更重要的是,如此薄的 APD 意味着偏压可以足够小但却能引发雪崩增益,同时器件速度也会很快.另外,集成将变得容易实现[40,41],如图 7.19 所示.在这个最新结构中,厚度 140 nm,宽度 750 nm 的 Ge 层沉积在 SiON 隔离层上,在隔离层下有约 100 nm 厚的波导.一系列的金属叉指沉积在 Ge 层上,间隔约 200 nm.因为 Ge 的折射率稍高于 Si,因此光能成功地通过倏逝耦合到 Ge 层.在非常靠近电极处,电场竟可以超过 120 kV/cm,如此强的电场将足够诱导产生碰撞离子化过程.

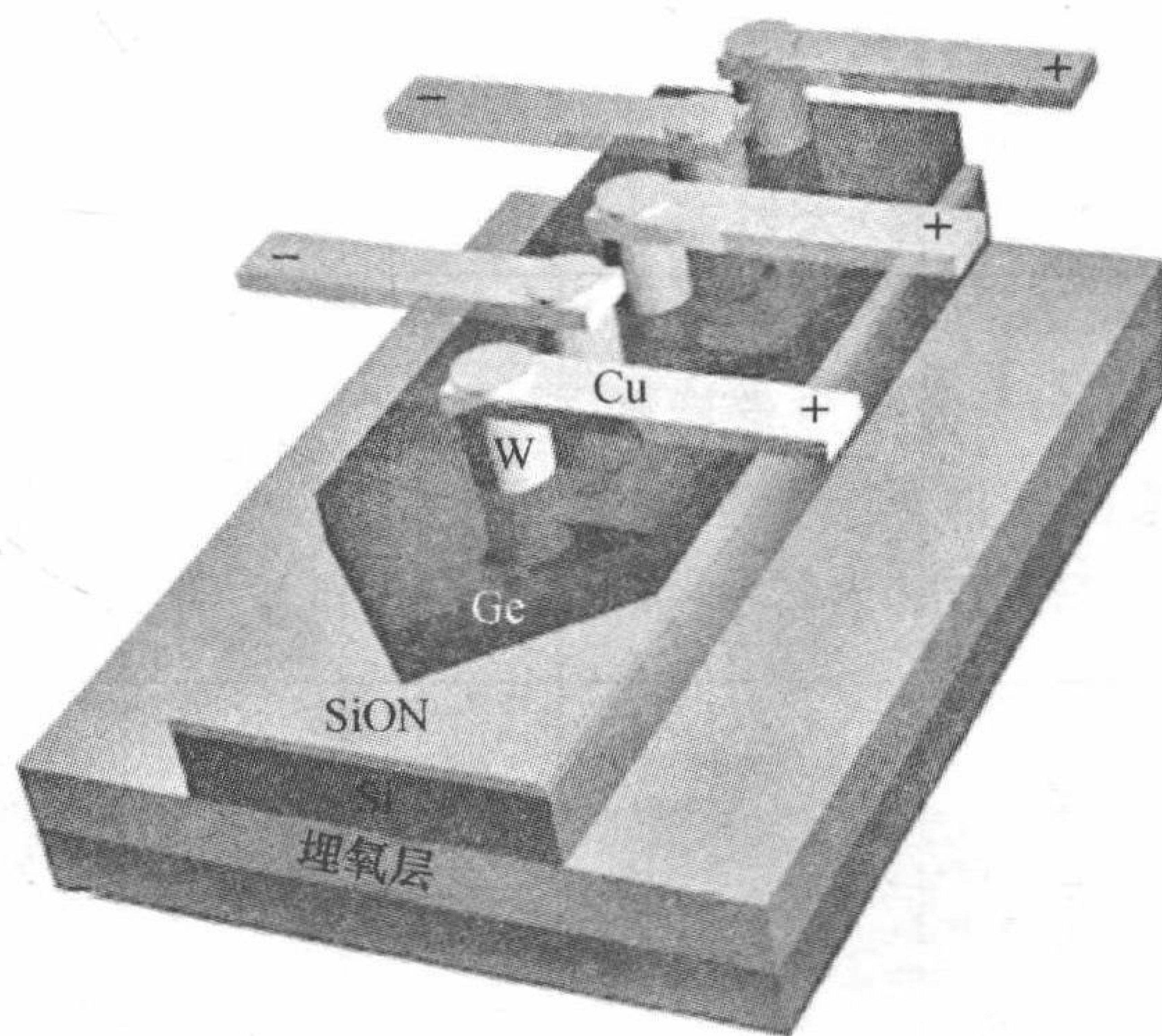


图 7.19 非一致性场 Ge APD 结构示意图

实验测试结果表明,在波长为 1.3 μm 和 1.5 μm 处的响应度分别为 0.4 A/W 和 0.14 A/W.当叉指间隔为 200 nm 时,随着偏压从 0.1 V 升至 1.1 V,3 dB 带宽

从 5 GHz 不断上升到 34 GHz，同时雪崩增益也在不断上升，如图 7.20 所示. 优异的结果显示了其巨大的潜力，不过，由于继承了倏逝耦合所带来的响应度低的问题，Ge APD 的工作还有一段路要走.

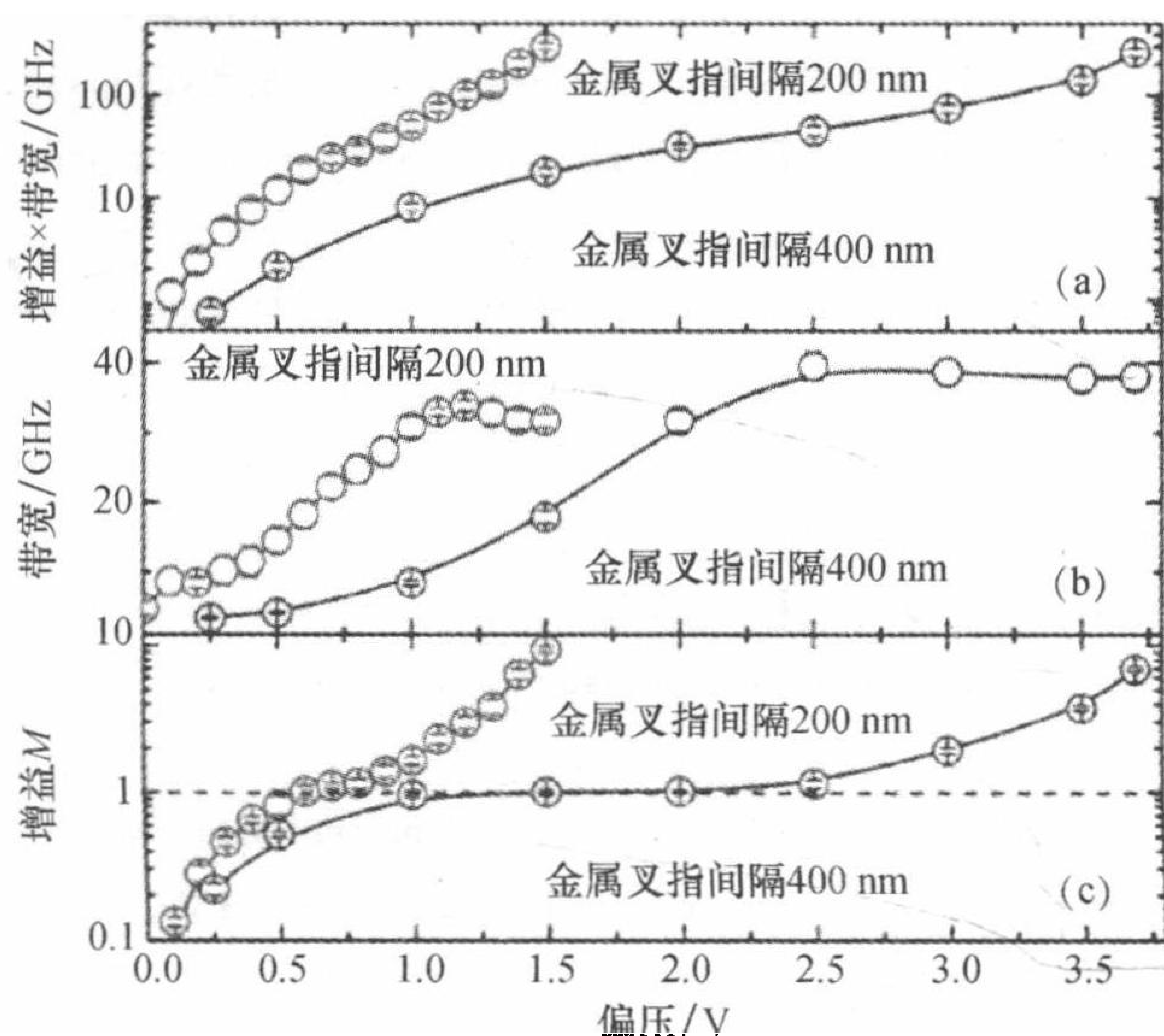


图 7.20　雪崩增益及带宽随偏压的变化

§7.5　总　　结

本章首先介绍光吸收三种主要方式：带间光吸收、自由载流子光吸收和双光子吸收，然后描述了三种基本的光电探测器的结构：p-i-n、APD 和 MSM 的工作原理和优缺点. 在此基础上介绍了两种探测器：硅光探测器和锗硅探测器，重点介绍了在近红外波段具有良好光响应特性的锗红外光探测器，并从 Ge/Si 材料的基本物理特性出发探讨了获得高质量锗薄膜的几种方法：外延技术、低温晶片键合及离子切割技术、快速融化生长技术. 最后主要介绍了最近几年国际上研究热点——波导型锗探测器的发展和最新的研究进展.

参考文献

[1] 余金中. 半导体光电子技术. 北京：化学工业出版社，2003.

[2] 张季熊. 光电子学教程. 广州：华南理工大学出版社，2009.

[3] Pavesi L，Lockwood D J. Silicon Photonics. Berlin：Springer，2004.

[4] 周志文. Si 基 SiGe-Ge 弛豫衬底生长及其 Ge 光电探测器研制[博士学位论文]. 厦门：厦

门大学,2009.

[5] 安毓英,曾晓东.光电探测原理. 西安：西安电子科技大学出版社,2004.

[6] 唐天同. 集成光电子学. 西安：西安交通大学，2005.

[7] 张宇. 光波导 GeSi 光电探测器的研究[硕士学位论文]. 天津：天津工业大学,2007.

[8] 成步文. 硅基高速 OEIC 光接收机芯片的研究. 光电子·激光,2003,14：599.

[9] Yang M, Rim K, Rogers D L. A high-speed, high-sensitivity silicon lateral tranch photodetector. IEEE Electr. Dev. Lett., 2002, 23: 395.

[10] Lee H C, Zeghbroeck B V. A novel high-speed silicon MSM photodetector operating at 830nm wavelength. IEEE Electr. Dev. Lett., 1995,16: 175.

[11] Levine B F ,Wynn J D,Klemens F P. 1Gb/s Si high quantum efficiency monolithically integrable $\lambda=0.88\ \mu m$ detector. Appl. Phys. Lett.,1995, 66(22): 2984.

[12] Emsley M K, Dosunmu O, Unlu M S. High-speed resonant-cavity enhanced silicon photodetectors on reflecting silicon on insulator substrates. IEEE Photo. Tech. Lett., 2002,14: 519.

[13] Ishikawa Y, et al. Strain-induced enhancement of near-infrared absorption in Ge epitaxial layers grown on Si substrate. J. Appl. Phys., 2005,98(1):3501.

[14] Luryi S,Kastalsky A, Bean J C. New infrared detector on a silicon chip. IEEE Trans. Electr. Devi., 1984,31(9): 1135.

[15] Samavedam S B, et al. High-quality germanium photodiodes integrated on silicon substrates using optimized relaxed graded buffers. Appl. Phys. Lett., 1998,73(15):2125.

[16] Tersoff J, LeGoues F K. Competing relaxation mechanisms in strained layers. Phys. Rev. Lett., 1994,72(22):3570.

[17] Tomp R M , Rosse F N, et al. Instability-driven SiGe island growth. Phys. Rev. Lett., 2000,84(20):4641.

[18] People R,Bean J C. Calculation of critical layer thickness versus lattice mismatch for Ge_xSi_{1-x}/Si strained-layer heterostructures. Appl. Phys. Lett.,1985,47:322.

[19] Ishikawa Y, et al. Strain-induced band gap shrinkage in Ge grown on Si substrate. Appl. Phys. Lett., 2003,82(13):2044.

[20] 聂辉文,成步文. 硅基锗材料的外延生长及其应用. 中国集成电路,2010.

Reed T. Silicon Photonics: the state of the art. West Sussex: John Wiley&Sons Ltd., 2008.

[21] 陈荔群,周志文,李成. Si 基 Ge 波导光电探测器的制备和特性研究. 光电子·激光, 2009,26(6): 1012.

[22] Thomas S G, Bharatan S, Jones R E, et al. Structural characterization of thick, high-quality epitaxial Ge on Si substrates grown by low-energy plasma-enhanced chemical vapor deposition. J. Electr. Mat., 2003,32(9):976.

[23] Luo G L, Yang T H, Chang E Y, et al. Growth of high-quality Ge epitaxial layers on Si (100). Jpn.J. Appl. Phys., 2003,42:L517.

[24] Luan H C, Lim D R, Lee K K, et al. High-quality Ge epilayers on Si with low threading-dislocation densities. Appl. Phys. Lett. ,1999,75(19):2909.

[25] Loh T H, Nguyen H S, Tung C H, et al. Ultrathin low temperature SiGe buffer for the growth of high quality Ge epilayer on Si(100) by ultrahigh vacuum chemical vapor deposition. Appl. Phys. Lett. , 2007,90(9):2108.

[26] Vanamu G, Datye S A K, Zaidi H. Epitaxial growth of high-quality Ge films on nanostructured silicon substrates. Appl. Phys. Lett. , 2006, 88(20):4104.

[27] Vivien L, Rouvière M,Fédéli J M,et al. High speed and high responsivity germanium photodetector integrated in a Silicon-On-Insulator microwaveguide. Opt. Exp. , 2007, 15: 9843.

[28] Chen L, Dong P, Lipson M. High performance germanium photodetectors integrated on submicron silicon waveguides by low temperature wafer bonding. Opt. Exp. 16,2008, 11513.

[29] Chen L, Lipson M. Ultra-low capacitance and high speed germanium photodetectors on silicon. Opt. Exp. , 2009,17: 7901.

[30] Fedelí J M, Migette M,Cioccio L D, et al. Incorporation of a photonic layer at the metallization levels of a CMOS circuit. Proc. IEEE International Conference on Group IV Photonics,2006.

[31] Liu Y C,Deal M D, Plummer J D. High-quality single-crystal Ge on insulator by liquid-phase epitaxy on Si substrates. Appl. Phys. Lett. , 2004,84(14): 2563.

[32] Ahn D, Hong C Y, Liu J, et al. High performance, waveguide integrated Ge photodetectors. Opt, Exp. , 2007, 15: 3916.

[33] Yin T, Cohen R, Morse M M, et al. 31GHz Ge n-i-p waveguide photodetectors on silicon-on-insulator substrate. Opt. Exp. , 2007, 15: 13965.

[34] Wang J, Loh W Y, Chua K T,et al. Evanescent-coupled Ge p-i-n photodetectors on Si-waveguide with seg-Ge and comparative study of lateral and vertical p-i-n configurations. IEEE Electr. Dev. Lett. , 2008, 29(5): 445.

[35] Wang J, Loh W Y, Chua K T, et al. Low-voltage high-speed (18 GHz/1V) evanescent -coupled thin-film-Ge Lateral p-i-n photodetectors integrated on Si waveguide. IEEE Photo. Tech. Lett. , 2008, 20(17): 1485.

[36] Vivien L, Osmond J, Fédéli J M, et al. 42GHz p-i-n. Germanium photodetector integrated in a silicon-on-insulator waveguide. Opt. Exp. , 2009,17: 6252.

[37] Feng D, et al. High-speed Ge photodetector monolithically integrated with large cross-section silicon-on-insulator waveguide. Appl. Phys. Lett. , 2009, 95(26): 1105.

[38] Liao S R, Feng N N,Feng D Z, et al. 36 GHz submicron silicon waveguide germanium photodetector. Opt. Exp. , 2011, 19(11): 10967.

[39] Colace L, Masini G, Galluzzi F, et al. Metal-semiconductor-metal near-infrared light

detector based on epitaxial Ge/Si. Appl. Phys. Lett., 1998, 72(24):3175.

[40] Ang K W, Zhu S Y, Yu M Bin, et al. High-performance waveguided Ge-on-SOI metal - semiconductor-metal photodetectors with novel silicon-carbon (Si:C) Schottky barrier enhancement layer. IEEE Photo. tech. Lett., 2008, 20(9): 754.

[41] Assefa S, Xia F N, Vlasov Y A, Reinventing germanium avalanche photodetector for nanophotonic on-chip optical interconnects. Nature, 2010, 464: 80.

第八章　硅基表面等离子激元器件

表面等离子激元(surface plasmon polariton,SPP)是在金属-电介质界面上存在的一种特殊的电磁表面波模式.由于其独特的表面波特性,它能够将光波约束在空间尺寸远小于其自由空间波长的区域.利用表面等离子激元对光波在亚波长尺寸上进行控制与操纵已经成为国际上的研究热点之一.除此之外,表面等离子激元还在光存储、光传感、超分辨成像及负折射率材料等许多领域有着广泛的应用前景.目前对表面等离子激元研究已成为一门科学,称为表面等离子学(plasmonics),是当前纳米光电子学的主要分支之一.

本章将首先概述表面等离子激元的基本概念和特性,然后分析金属-电介质界面的表面等离子激元的局域增强特性,最后介绍一些主要的有源和无源表面等离子激元器件.

§8.1　表面等离子激元概述

表面等离子激元的研究最早可以追溯到 19 世纪末.当时索末菲与 Zenneck 最早给出了在射频波段沿着有限电导率导体的表面上传输的表面波的数学描述.1902 年,Wood[1] 首先观察到在可见光波段当光入射到金属光栅时其反射光谱存在反常衍射现象,并且对这种现象进行了公开描述.1941 年,Fano[2] 用金属与空气界面的表面电磁波激发模型对这一现象给出了解释.将金属光栅的反常衍射现象与之前索末菲的理论工作联系起来,提出了表面等离子波的概念.1957 年,Ritchie[3] 发现,当电子穿过金属薄片时存在一定数量的消失峰.他将这种消失峰称之为"能量降低的"等离子模式,并指出了这种模式和薄膜边界的关系,第一次提出了用于描述金属内部电子密度纵向波动的"金属等离子体"概念.两年后,Powell 和 Swan 用实验证实了 Ritche 的理论.1968 年,Kretschmann 等通过棱镜耦合的办法实现了可见光波段的索末菲表面波[4],由此关于以上所有这些现象的统一理论解释—表面等离子激元理论就建立起来了.表面等离子激元是由横磁(TM)波和金属表面电子的相互耦合振动而形成的,如图 8.1 所示.与自由空间或者电介质材料中的电磁场相比,表面等离子激元在其偏振状态、模式体积、色散特性等方面具有若干特殊的物理性质,为新型光电功能器件及光电一体芯片集成技术的发展提供了一条新的途径.

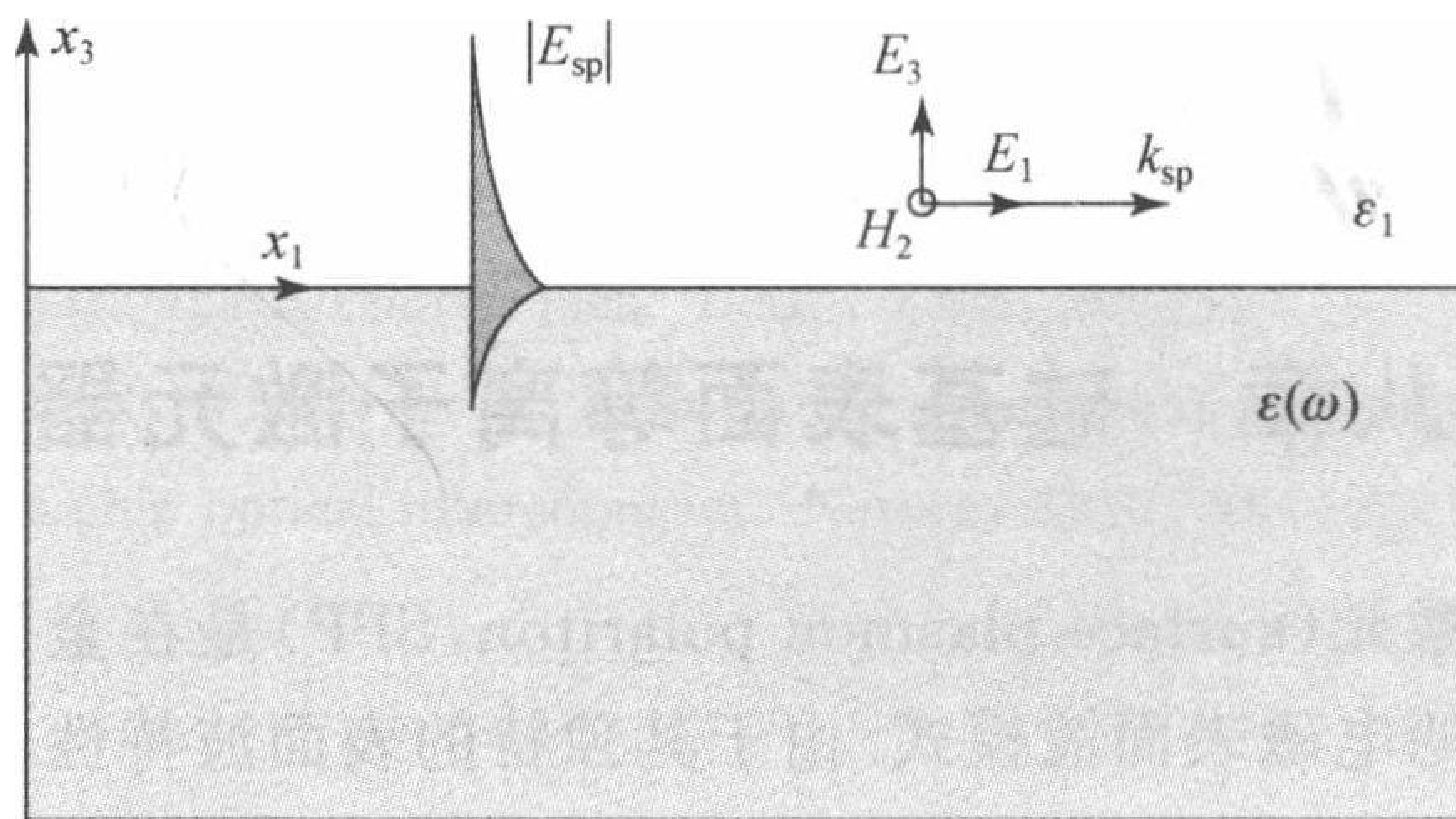


图 8.1　金属-电介质界面中的表面等离子激元

8.1.1　金属的电磁特性

由于金属具有的唯一特性就是它能够维持一种由光感应出来的金属表面自由电子的集体共振激发，即表面等离子体激元. 所以，金属的介质常数及其色散模型的讨论显得尤为重要. 能够产生表面等离子激元所用到的金属几乎都是一些贵重金属，金属与电磁场之间的相互作用仍可以在经典电磁理论的框架内通过麦克斯韦方程来解释.

金属中的电磁特性可以通过其复介质常数来描述，并且在相当宽的频域范围内，可以通过德鲁德模型来解释[5]. 金属德鲁德模型是自由电子气体模型. 它假设除了在金属表面层受到电势作用外，传导电子是完全自由的. 表面电势的作用是将传导电子限制在样品内部. 传导电子如同理想气体中的分子-自由电子气体. 在此模式中，密度为 n 的自由电子在正离子的背景中运动，且其晶格势及电子之间相互作用的细节并不考虑. 相反，人们假定晶格结构对电子运动的影响被包含进电子的有效质量 m^* 中. 电子在外界电磁场的作用下产生振荡，电子之间的相互作用是通过碰撞频率 γ 来描述，$\gamma=1/\tau$，τ 为自由电子气弛豫时间，在室温下通常为 10^{-14} s 数量级.

自由电子气中的电子在外界电场 $\boldsymbol{E}$ 的作用下的运动方程可以由下式表示：

$$m\frac{\mathrm{d}\boldsymbol{x}^2}{\mathrm{d}t^2}+m\gamma\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}=-e\boldsymbol{E}, \tag{8.1.1}$$

式中 $\boldsymbol{x}$ 为电子位移，e 为电子电量. 假设电场 $\boldsymbol{E}$ 是以时谐波 $E=x_0\mathrm{e}^{-\mathrm{i}\omega t}$ 的形式存在，那么式(8.1.1)的一个描述电子振动的特别解为 $x(t)=x_0\mathrm{e}^{-\mathrm{i}\omega t}$. 将此特别解代入式(8.1.1)中可以得到

$$\boldsymbol{x}(t)=\frac{e}{m(\omega^2+\mathrm{i}\gamma\omega)}\boldsymbol{E}(t). \tag{8.1.2}$$

从而，自由电子运动对电极化强度 $\boldsymbol{P}$ 带来的贡献 $\boldsymbol{P}=-ne\boldsymbol{x}$ 可以精确表示为

$$P = -\frac{ne^2}{m(\omega^2 + i\gamma\omega)}E, \tag{8.1.3}$$

再根据电位移矢量 $\boldsymbol{D}$ 与 $\boldsymbol{P}$ 的关系 $\boldsymbol{D}=\varepsilon_0\boldsymbol{E}+\boldsymbol{P}$(其中 ε_0 为真空中的介电常数)，可以得到

$$D = \varepsilon_0\left(1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}\right)E, \tag{8.1.4}$$

式中 $\omega_p^2=\frac{ne^2}{(\varepsilon_0 m)}$，$\omega_p$ 称为金属的等离子体频率.

到此，金属的相对介电常数 $\varepsilon_m(\omega)$ 为

$$\varepsilon_m(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + \gamma^2} + i\frac{\gamma}{\omega}\frac{\omega_p^2}{\omega^2 + \gamma^2}. \tag{8.1.5}$$

金属的相对介电常数是一个复数，即 $\varepsilon_m=\varepsilon_{rm}+i\varepsilon_{im}$，其中 ε_{rm} 和 ε_{im} 分别是 ε_m 的实部和虚部，阻尼系数由方程 $\gamma=q^2\omega^2/(6\pi\varepsilon_0 mc^3)$ 决定. 将电子电量 q、电子质量 m、光速 c 和光的圆频率 $\omega=10^{15}$ rad/s 等物理常数代入以上方程式，可得到 $\gamma\approx10^7$ Hz. 在可见光与近红外光波段，金属的相对介电常数的实部是一个绝对值较大的负数，虚部是一个较小的正数，即 $\varepsilon_{rm}<0$，$|\varepsilon_{rm}|\gg\varepsilon_{im}$. 图 8.2 所示的是银、金、铜、铝、镍、铬和钨等几种常见的贵金属的 $\varepsilon_m(\omega)$ 的 ε_{rm} 和 ε_{im} 的大小与光波长的关系.

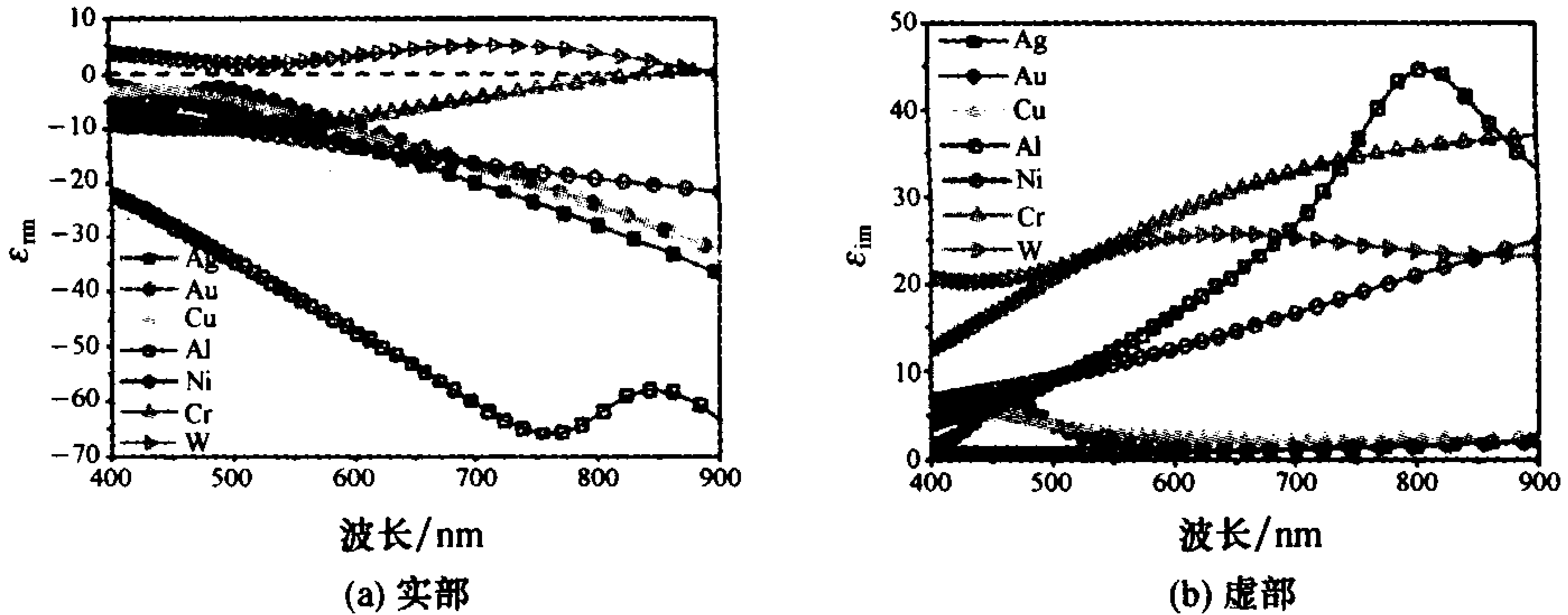


图 8.2　银、金、铜、铝、镍、铬和钨等七种金属的相对介电常数与光波长的关系

另外，金属的相对介电常数可以简化为

$$\varepsilon_m(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}. \tag{8.1.6}$$

金属的相对介电常数的实部 $\varepsilon_{rm}(\omega)$ 与虚部 $\varepsilon_{im}(\omega)$ 分别由下面两式给出：

$$\varepsilon_{rm}(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + \gamma^2}, \tag{8.1.7}$$

$$\varepsilon_{im}(\omega) = \frac{\omega_p^2\gamma}{\omega(\gamma^2 + \omega^2)}. \tag{8.1.8}$$

下面对金属在不同频率下的电磁特性做简单的讨论. 由式(8.1.6)可知，在电磁波频率 $\omega \gg \gamma$ 的条件下，式(8.1.7)中的 $\varepsilon_{\rm m}$ 可近似为 $1-\omega_{\rm p}^2/\omega^2$. 由此可以看出，若 $\omega < \omega_{\rm p}$，则金属介电常数实部为一负值，并且虚部 $\varepsilon_{\rm im}(\omega) \approx \gamma\omega_{\rm p}^2/\omega^3 \ll |\varepsilon_{\rm rm}|$. 对于金、银和铜等金属而言，其 $\omega_{\rm p}$ 值都位于紫外光频率范围，且 γ 远小于可见光频率，所以这些金属的介电常数在可见光频率范围内皆可符合上述条件. 譬如银，在可见光和红外光波段内，等离子体频率 $\omega_{\rm p} \approx 1.374 \times 10^{16}$ rad/s，阻尼系数 $\gamma \approx 3.21 \times 10^{13}$ rad/s. 当 $\omega \ll \omega_{\rm p}$ 时，其介电常数的实部通常是一个绝对值很大的负数，可以认为是负无穷大；而虚部则很小，可以忽略不计，因而银可以被看做理想的具有表面等离子体效应的金属.

在光波段，物质的光学性质通常由其折射率来描述. 金属的折射率仍然为复数，定义为 $\tilde{n} = \eta + {\rm i}\kappa = \sqrt{\varepsilon(\omega)}$，则 η 与 κ 分别为

$$\begin{cases} \eta = \left[\dfrac{1}{2}\left(\sqrt{\varepsilon_1^2 + \varepsilon_2^2} + \varepsilon_1\right)\right]^{1/2}, \\ \kappa = \left[\dfrac{1}{2}\left(\sqrt{\varepsilon_1^2 + \varepsilon_2^2} - \varepsilon_1\right)\right]^{1/2}. \end{cases} \tag{8.1.9}$$

当 ω 在小于 $\omega_{\rm p}$ 且大于 γ 的范围内时，$\kappa \gg \eta$. 由于电磁波随空间与时间变化正比于 ${\rm e}^{{\rm i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = {\rm e}^{-\boldsymbol{k}_{\rm I}\cdot\boldsymbol{r}}{\rm e}^{{\rm i}(\boldsymbol{k}_{\rm R}\cdot\boldsymbol{r}-\omega t)}$，其中 $k_{\rm R}$ 与 $k_{\rm I}$ 分别表示波向量的实部与虚部，其大小分别表示为 $k_{\rm R} = \eta\omega/c$ 与 $k_{\rm I} = \kappa\omega/c$，所以对于此波段电磁波而言，其在金属内部之传播性质主要是由波矢向量虚部所主导，亦即电磁场振幅或能量将会很快地随着传播距离呈指数衰减而无法深入穿透至金属内部若定义金属趋肤深度(skin depth，δ)为电磁波振幅衰减至原本振幅的 1/e 时的穿透距离，则 $\delta = 1/k_{\rm I} = c/(\omega\kappa)$，其值随外加电磁波频率变化而变化. 此外，从式(8.1.6)也可以看出，当 $\omega > \omega_{\rm p}$ 之后，金属介电常数或复数折射率的实部为小于 1 的正数，且虚部将趋近于零，在此频率范围的电磁波将可以穿透金属内部而传播.

为了描述电子带间跃迁对金属相对介电常数带来的影响，人们对德鲁德模型做了一定修正，用 ε_∞ 来描述当 $\omega \gg \omega_{\rm p}$ 时的 $\varepsilon(\omega)$ 值，这个值也就是电子带间跃迁引起的介电常数值. 这样，金属的相对介电常数可以修正为

$$\varepsilon_{\rm m}(\omega) = \varepsilon_\infty - \frac{\omega_{\rm p}^2}{\omega^2 + {\rm i}\gamma\omega}. \tag{8.1.10}$$

式(8.1.10)能更好地描述金属的实际情况，从而被广泛应用于表面等离子激元研究中对金属的描述. 除了能对金属的介电常数作出理论解释外，德鲁德模型的另一个重要的优点就是它可以被引入到时域的电磁计算方法中，比如时域有限差分(FDTD)法，用于在较宽频谱范围内对材料物理参数的描述，从而可以通过分析结构对脉冲的响应而快速得到频域上宽频内的结果.

8.1.2　金属-电介质界面表面等离子激元

如同光波在量子力学中被称为光子(photon)、晶格振动在固体物理学中被量子化描述成声子(phonon)一样，金属表面中自由电子气密度的涨落被量子化为表面等离子激元. 表面等离子激元是一种存在于不同的介电常数构成的材料表面或者界面从而导致与表面有关的特殊激励. 具有正的介电常数 ε_d 的电介质材料如二氧化硅、硅等和具有负的介电常数 ε_m 材料，如金、银、铜和镍等贵金属，能够产生一种被限制在界面处附近传播的电磁波. 如果电磁波对金属表面的自由电子密度波产生激励，且又被耦合，那么，它就是表面等离子激元[6~10].

表面等离子激元是光在金属-电介质界面中感应出的电磁波，电磁波是物质存在的一种形式，作为波，它携带着能量，以一定的速度在空间传播. 一般来说，在金属-电介质界面中，金属是指金、银、铜和镍等能够支持与激发表面等离子激元的贵金属；电介质一般包括空气、二氧化硅和硅等材料. 如果在金属-电介质界面中传播的电磁波电矢量 $\boldsymbol{E}$ 垂直于金属表面的分量不为零，那么金属表面的自由电子密度就会发生集体起伏和振荡形成电荷密度波，从而感生出表面等离子激元，如图 8.3 所示. 表面等离子激元平行于金属-电介质界面方向是传播的，而在垂直于金属-电介质界面方向是衰减的. 在垂直于界面的方向上呈指数衰减，而在界面的附近局域场强非常大. 金属中的自由电子在外界电磁场的作用下相对金属中的正离子发生相对位移，带来电子密度的重新分布. 从而在金属表面的两边产生电场 $\boldsymbol{E}_m$，$\boldsymbol{E}_d$，其大小分别为

$$\boldsymbol{E}_m(x,z,t)=\boldsymbol{E}_{m,0}e^{i(k_x x-k_z z-\omega t)},\quad \boldsymbol{E}_d(x,z,t)=\boldsymbol{E}_{d,0}e^{i(k_x x+k_z z-\omega t)}.$$

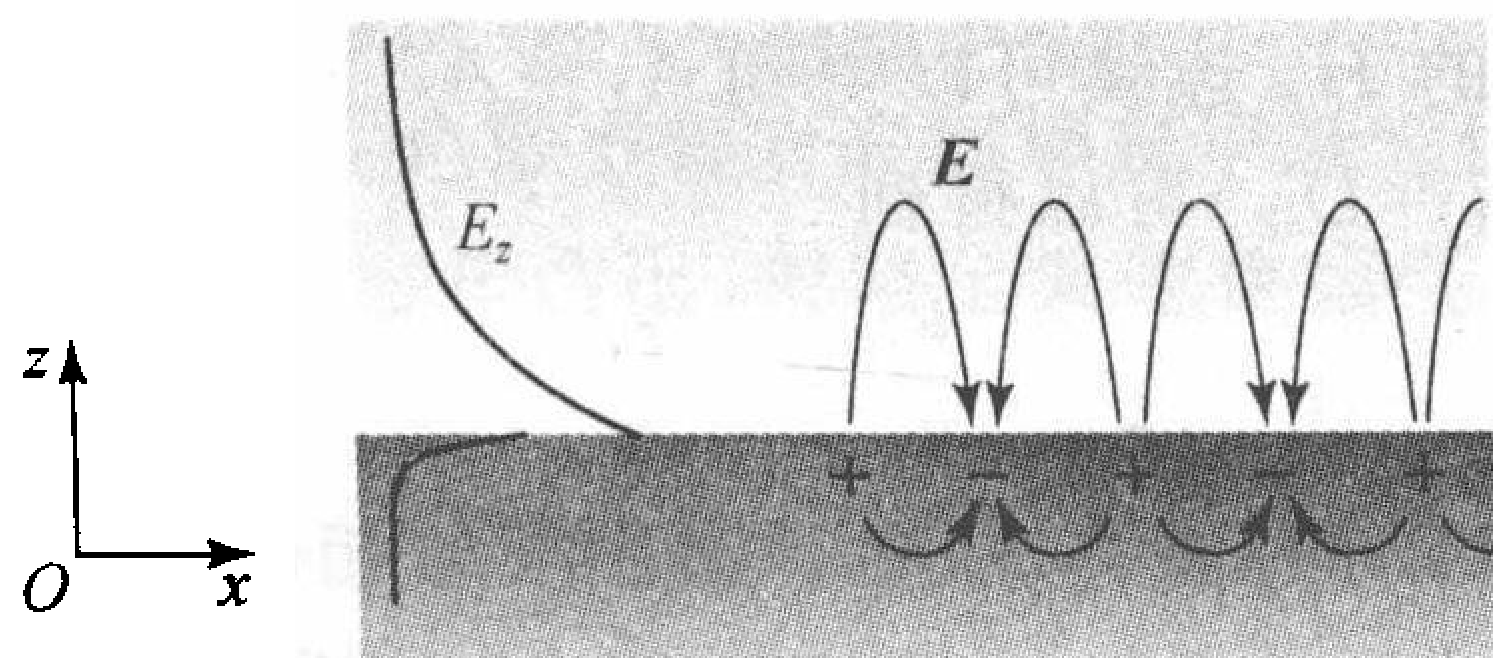


图 8.3　金属-电介质界面中的表面等离子激元电场与电子密度波状态分布

8.1.3　表面等离子激元色散关系

从电磁波的基本方程着手，可以推导出金属-电介质界面上表面等离子激元色散关系. 如图 8.4 所示的结构，$z>0$ 的半空间为介电常数为 ε_2 的介质；而 $z<0$ 的半空间为金属，介电常数为 $\varepsilon_1(\omega)$. 假设表面等离子激元沿着 x 方向传播.

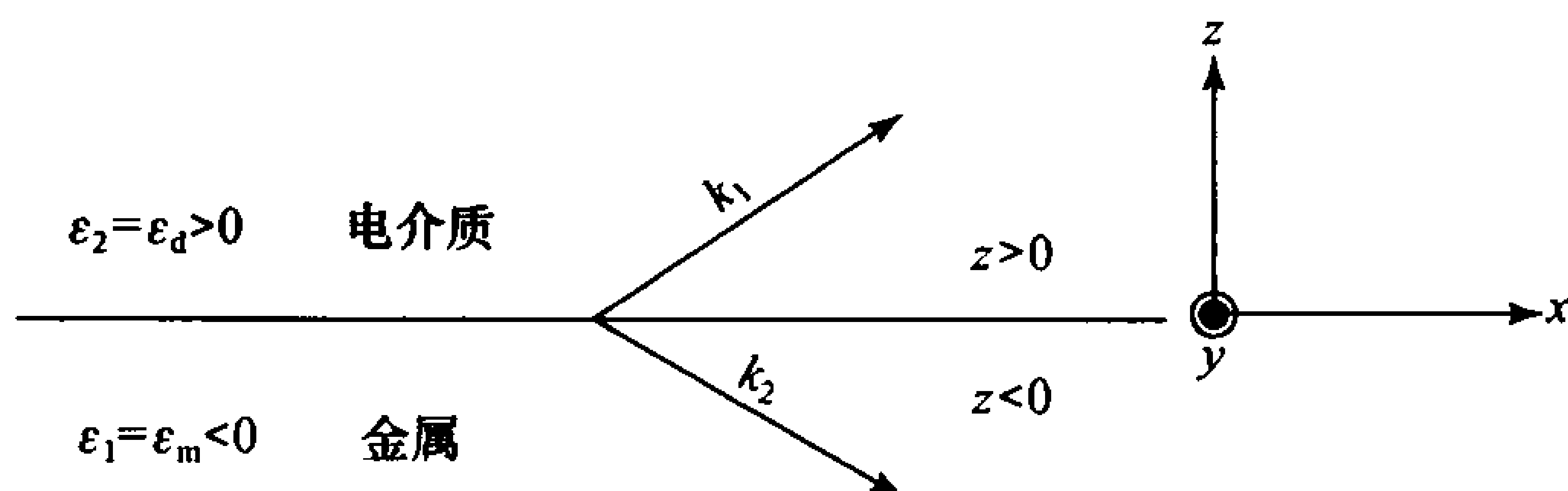


图 8.4 金属-电介质界面结构

首先考虑 TM 偏振,磁场沿着 y 方向分布. 考虑到表面等离子激元在 $z>0$ 及 $z<0$ 分别呈指数分布,即 $H_y=A_2\mathrm{e}^{\mathrm{i}\beta x}\mathrm{e}^{-k_2 z}(z>0)$, $H_y=A_1\mathrm{e}^{\mathrm{i}\beta x}\mathrm{e}^{k_1 z}(z<0)$,其中 $k_i(i=1.2)$为波矢在 z 方向上在两种介质中的分量,则在 $z>0$ 的半空间的场分布为

$$\begin{cases} H_y(z)=A_2\mathrm{e}^{\mathrm{i}\beta x}\mathrm{e}^{-k_2 z}, \\ E_x(z)=-\mathrm{i}\dfrac{1}{\omega\varepsilon_o\varepsilon_2}\dfrac{\partial H_y}{\partial z}=\mathrm{i}A_2\dfrac{1}{\omega\varepsilon_o\varepsilon_2}k_2\mathrm{e}^{\mathrm{i}\beta x}\mathrm{e}^{-k_2 z}, \\ E_z(z)=-\dfrac{\beta}{\omega\varepsilon_o\varepsilon_2}H_y=-A_2\dfrac{\beta}{\omega\varepsilon_o\varepsilon_2}\mathrm{e}^{\mathrm{i}\beta x}\mathrm{e}^{-k_2 z}; \end{cases} \tag{8.1.11}$$

而在 $z<0$ 的半空间中,同样可以得到各个场的量分别为

$$\begin{cases} H_y(z)=A_1\mathrm{e}^{\mathrm{i}\beta x}\mathrm{e}^{k_1 z}, \\ E_x(z)=-\mathrm{i}A_1\dfrac{1}{\omega\varepsilon_o\varepsilon_1}k_1\mathrm{e}^{\mathrm{i}\beta x}\mathrm{e}^{k_1 z}, \\ E_z(z)==-A_1\dfrac{\beta}{\omega\varepsilon_o\varepsilon_1}\mathrm{e}^{\mathrm{i}\beta x}\mathrm{e}^{k_1 z}. \end{cases} \tag{8.1.12}$$

根据在界面 $z=0$ 处的 H_y 连续可以得到 $A_1=A_2$. 再根据 $z=0$ 处电位矢量在 z 方向的分量连续($D_{z1}=D_{z2}$)可以得到

$$\frac{k_2}{k_1}=-\frac{\varepsilon_2}{\varepsilon_1}, \tag{8.1.13}$$

由于 $k_1, k_2>0$,式(8.1.13)说明 ε_1 与 ε_2 必须符号相反.

由此可见,表面等离子激元只存在于相对介电常数符号相反的两种材料之间的界面,而在光波段金属与电介质之间的界面正好满足这种条件. 根据金属与电介质中的波矢关系,可以得到

$$\begin{cases} \beta^2+(\mathrm{i}k_1)^2=\varepsilon_1 k_0^2, \\ \beta^2+(\mathrm{i}k_2)^2=\varepsilon_2 k_0^2. \end{cases} \tag{8.1.14}$$

将式(8.1.14)代入到式(8.1.13)中可以得到金属-电介质界面中表面等离子激元曲线为

$$\beta=\sqrt{\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1+\varepsilon_2}}. \tag{8.1.15}$$

由于 ε_1, ε_2 符号相反，要使得式(8.1.15)有解，必须满足 $\varepsilon_1+\varepsilon_2<0$，即 $\varepsilon_1(\omega)<-\varepsilon_2$. 也就是说，只有金属的相对介电常数实部为负且绝对值大于电介质的相对介电常数时，在金属-电介质的界面上才有表面等离子激元的存在.

接下来再分析 TE 偏振的情况，假设 TE 偏振的表面波形式存在，电场沿着 y 向分布，则在 $z>0$ 区域的场分布情况为

$$\begin{cases} E_y(z)=A_2 e^{i\beta x}e^{-k_2 z}, \\ H_x(z)=-iA_2\dfrac{1}{\omega\mu_o}k_2 e^{i\beta x}e^{-k_2 z}, \\ H_z(z)==A_2\dfrac{\beta}{\omega\mu_o}e^{i\beta x}e^{-k_2 z}; \end{cases} \tag{8.1.16}$$

在 $z<0$ 的场分布为

$$\begin{cases} E_y(z)=A_1 e^{i\beta x}e^{k_1 z}, \\ H_x(z)=iA_1\dfrac{1}{\omega\mu_o}k_1 e^{i\beta x}e^{k_1 z}, \\ H_z(z)=A_1\dfrac{\beta}{\omega\mu_o}e^{i\beta x}e^{k_1 z}. \end{cases} \tag{8.1.17}$$

由 $z=0$ 处的电场 E_y 及磁场分量 H_x 连续可以得到 $A_1(k_1+k_2)=0$. 由于假设 TE 偏振时同样是表面波的形式存在，这就说明 k_1, $k_2>0$. 因此上面这个条件只有 $A_1=0$ 才能够满足，然后可以推得 $A_1=A_2=0$. 这说明 TE 偏振的表面波形式并不存在. 从以上分析来看，金属与电介质之间的表面波形式只能为 TM 偏振态，这个表面波就是表面等离子激元. 表面等离子激元只能以 TM 偏振的形式存在是它的一个重要性质. 导致这种单偏振特性的根本原因在于金属在光波段范围内的电学性质与磁学性质，即 $\varepsilon<0$，而 $\mu>0$.

下面根据表面等离子激元色散方程对表面等离子激元性质进行分析：为了便于分析，这里先假设金属为理想情况，忽略金属德鲁德模型中的电子碰撞，即设 $\gamma=0$，则金属的相对介电常数为 $\varepsilon_1(\omega)=1-\omega_p^2/\omega^2$. 可见当 $\omega<\omega_p$ 时，金属相对介电常数为负. 由上面对式(8.1.15)的分析可知，表面等离子激元存在的条件是 $\varepsilon_1(\omega)<-\varepsilon_2$. 定义金属与电介质的相对介电常数互为相反数时的频率为 ω_{sp}，称为表面等离子体频率. 由 $\varepsilon_1(\omega)=-\varepsilon_2$ 得到

$$\omega_{sp}=\frac{\omega_p}{\sqrt{1+\varepsilon_2}}. \tag{8.1.18}$$

当 $\omega<\omega_{sp}$ 时，表面等离子激元存在于金属与电介质的界面上，这时其色散关系如图 8.5 中曲线所示，该图中的直线为介质中光的色散曲线. 从图中可见，当光频率比较小时，表面等离子激元的波矢与在电介质中光波的波矢大小比较接近；随着光频率的增大，两者之间的差距也逐渐拉大. 图 8.5 结果表明，表面等离子激元的波矢大于电介质中的光波矢. 因此，激发表面等离子激元需要光栅或者

棱镜耦合结构来辅助，从而满足波矢匹配条件.

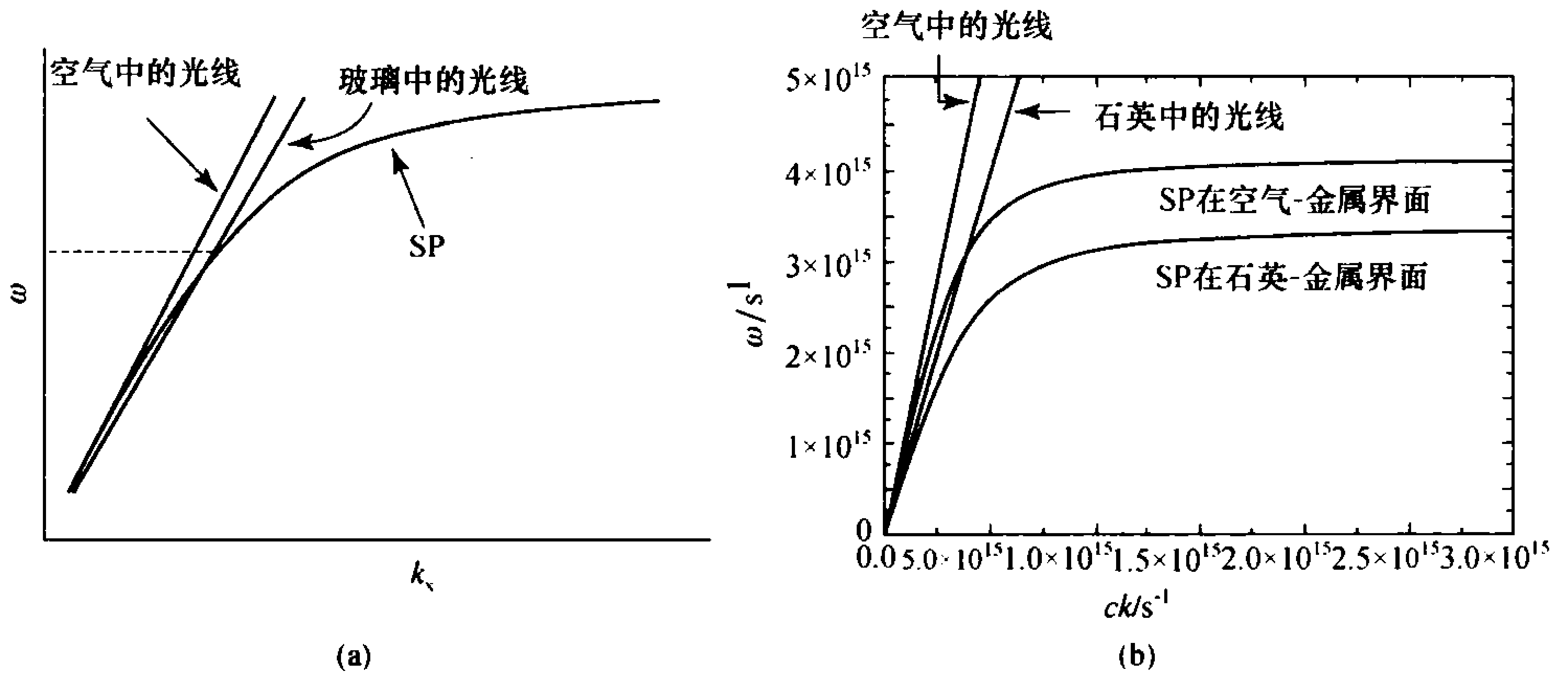


图 8.5 金属-电介质界面中表面等离子激元色散曲线(a)；金属-空气、金属-石英界面中的表面等离子激元色散曲线及在空气、石英中光线(b)

当 $\omega>\omega_p$ 时，金属表现为电介质特性，光能透过金属继续传播. 但由于金属强色散特性，光在其中的传播仍然与普通的电介质不同. 由 $\omega^2\varepsilon=c^2k^2$ 及 $\varepsilon=1-\omega_p^2/\omega^2$ 可以得到此时的色散方程为

$$\omega=\sqrt{\omega_p^2+c^2k^2}. \tag{8.1.19}$$

8.1.4 表面等离子激元激发

下面简单地讨论一下如何激发表面等离子激元，且激发条件又是什么？这主要取决于表面等离子激元色散关系. 根据，$k_{spp}=k_0\sqrt{\varepsilon_m\varepsilon_1/(\varepsilon_m+\varepsilon_1)}$，在光频区，金属的 $\varepsilon_m<0$，$|\varepsilon_m|\gg1$，有 $|\varepsilon_m+\varepsilon_1|<|\varepsilon_m|$，故 $k_{spp}>k_0$. 可以看到，表面等离子激元波矢不等于相同频率的光波矢. 这是因为表面等离子激元是光波与可迁移的表面电荷(如金属中自由电子)之间相互作用产生的电磁模，这个电磁模有大于同一频率下光子在真空中或周边电介质中的波数，这可以解析为什么在光滑的金属表面中光与表面等离子激元是不会相互耦合的原因. 换句话来说，表面等离子激元不能够被激发. 只有采用特殊的手段，例如，外加光栅耦合作用，金属光栅结构能够提供一个光栅动量 G 来满足其动量守恒，激发表面等离子激元[如图 8.6(a)所示]. 也可以从图 8.6(b)来看，光在光滑的金属表面不能感应出表面等离子激元，这是因为除了 $k=0$ 外，入射光和表面等离子激元的色散曲线没有相交点.

一般采用以下两种常用方法在金属-电介质界面中激发表面等离子激元：

① 棱镜耦合，可以使用电介质棱镜使入射光和表面等离子激元在棱镜和金属界面之间发生耦合作用.

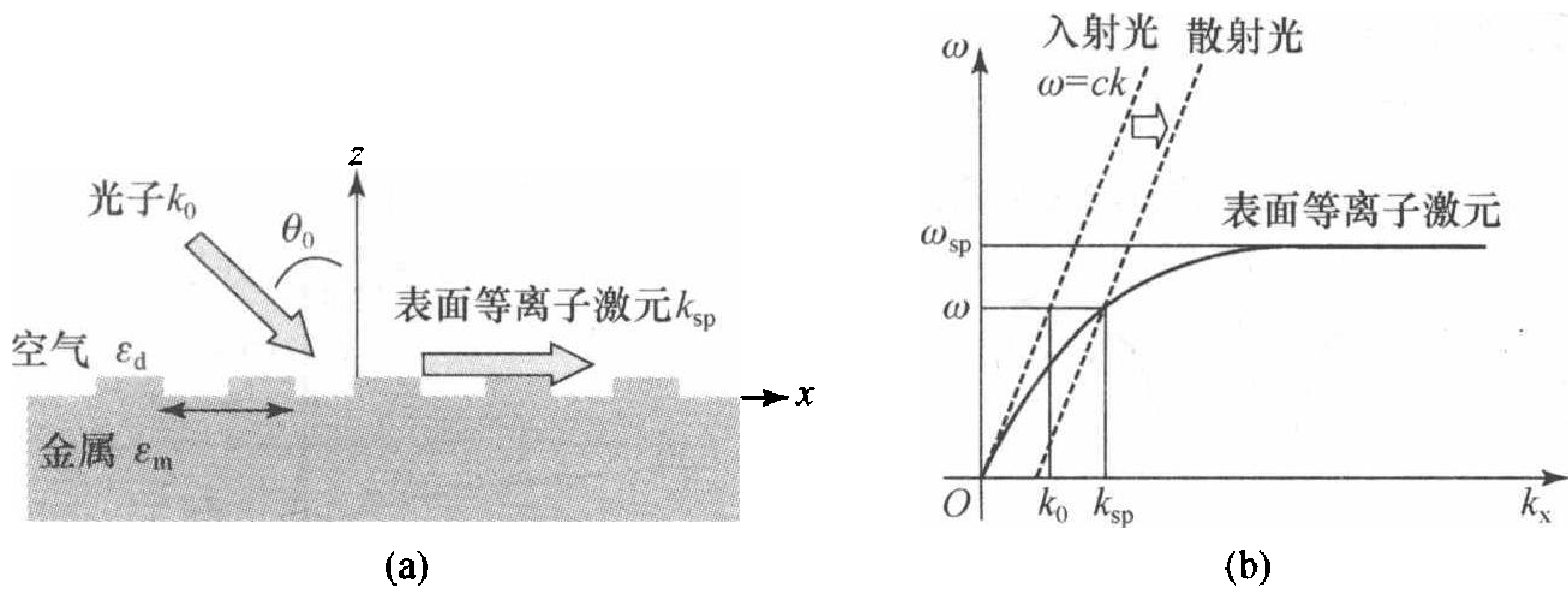


图 8.6　光在金属光栅耦合作用下激发表面等离子激元(a)；入射光、散射光和表面等离子激元色散曲线(b)

② 衍射光栅耦合，还可以在金属表面使用周期性光栅结构. 当光入射到这带有光栅结构的金属表面，入射光会被这种光栅结构所散射. 散射光的表面成分会从周期性的光栅结构中得到一个额外的动量. 这种额外的动量能够使散射光的表面成分有助于金属表面等离子激元的激发.

除棱镜与衍射光栅耦合外，表面等离子激元还有多种激发方式，如采用强聚焦光束激发、采用近场光学激发、利用带电粒子碰撞激发、利用波导边界处的倏逝波激发等.

8.1.5　表面等离子激元传播

表面等离子激元是沿着金属-电介质界面水平方向传播，而在垂直于其界面的方向呈指数衰减. 其传播的四个特征长度是重要的参量，也是在纳米等离子体光学器件设计中的重要参考指标. 它们分别是表面等离子激元传播长度 δ_{spp}，表面等离子激元波长 λ_{spp} 以及表面等离子激元场穿透进入电介质和金属中的趋肤深度 δ_d 和 δ_m. 这四个不同的特征长度数量级大小如图 8.7 所示[11,12].

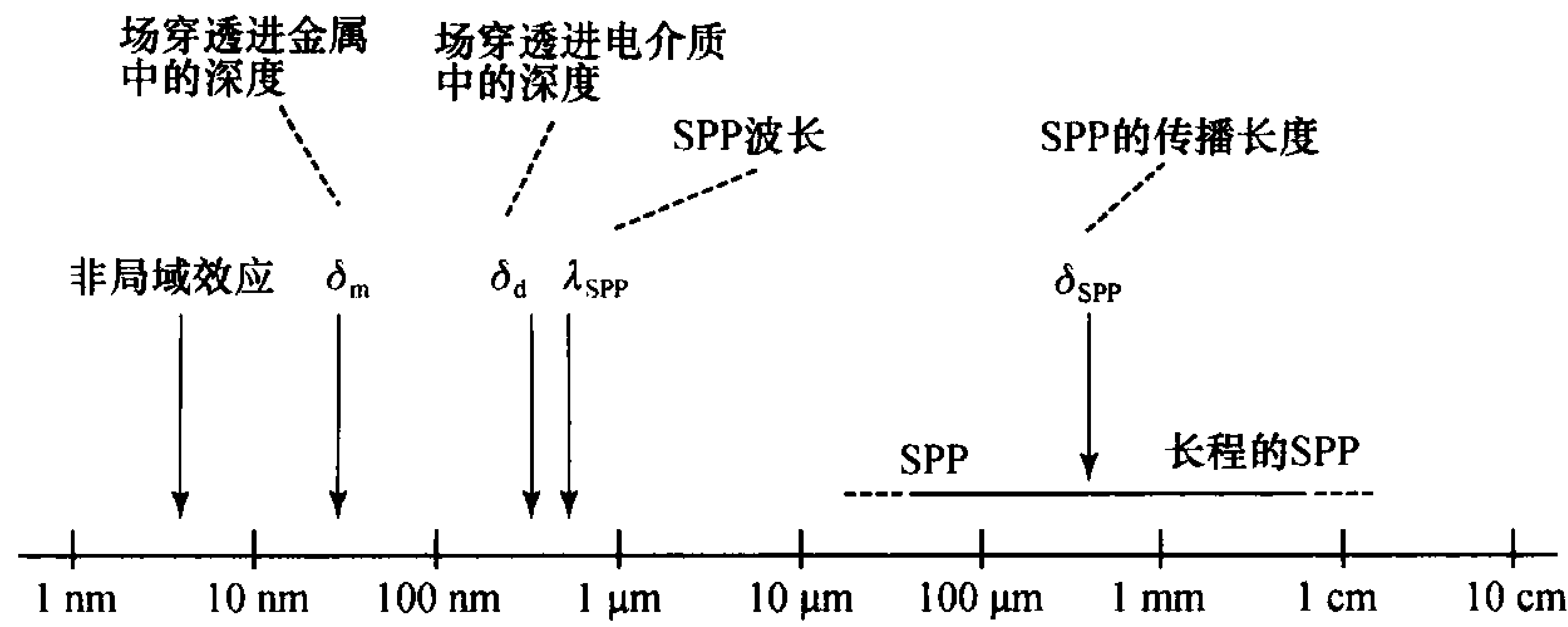
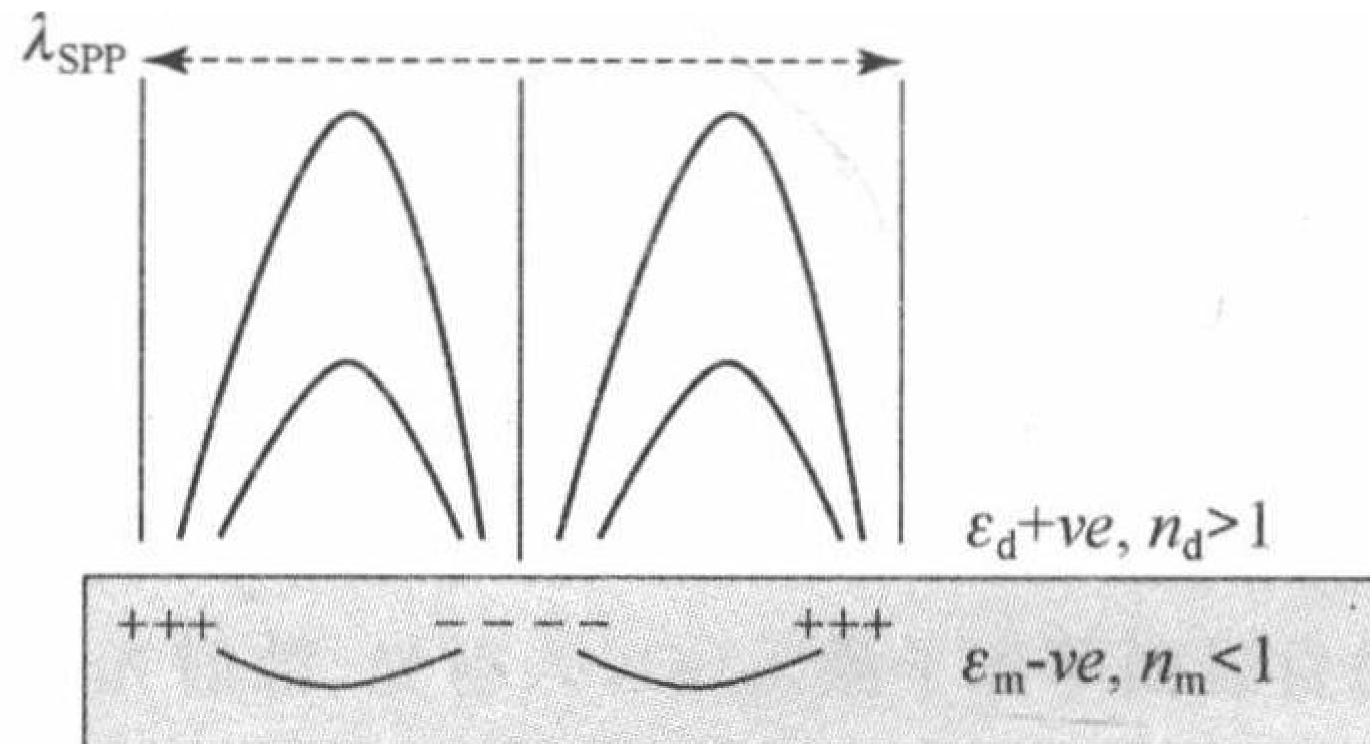
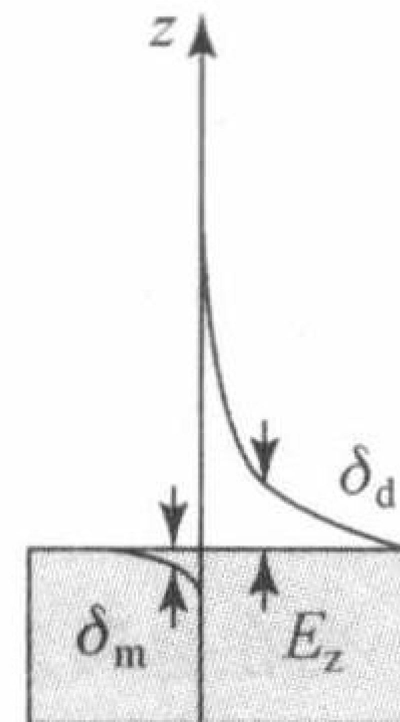


图 8.7　表面等离子激元在金属-电介质界面中传播的四个特征长度

如图 8.8(a)所示，表面等离子激元是沿着金属表面传播，其波长为 λ_{SPP}，能够传播的距离为 δ_{SPP}. 而图 8.8(b)所示的是表面等离子激元在垂直金属-电介质界面中的穿透趋肤深度分别为 δ_m 和 δ_d.



(a) 在金属-电介界面中被激发及表面电荷、电场分布



(b) 垂直入射到金属-电介质中的趋肤深度(δ_m和δ_d)

图 8.8　表面等离子激元

下面分别阐述这四个特征长度的定义及其表达式：

① 表面等离子激元波长 λ_{SPP}.

金属的介电常数 $\varepsilon_m(\omega)$是一个与 ω 有关的函数. 它是一个复数，即 $\varepsilon_m=\varepsilon_m'+i\varepsilon_m''$，其中 ε_m' 和 ε_m' 分别是其实部和虚部. 同样，波矢 k_{SPP}也是一个复数，即 $k_{SPP}=k_{SPP}'+ik_{SPP}'$. 从表面等离子激元波矢的实部 $k_{SPP}'=k_0\sqrt{\frac{\varepsilon_d\varepsilon_m'}{\varepsilon_d+\varepsilon_m'}}$ 可计算表面等离子激元波长

$$\lambda_{SPP}=\frac{2\pi}{k_{SPP}'}=\lambda_0\sqrt{\frac{\varepsilon_d+\varepsilon_m'}{\varepsilon_d\varepsilon_m'}}. \tag{8.1.20}$$

由此可见，λ_{spp}总是稍小于真空中光波长 λ_0. 如果在金属表面上加工成各种周期性调制结构(布拉格散射体)则可以实现对表面等离子激元的控制，那么，这个结构的周期必须与 λ_{SPP}同一数量级，或者几倍于 λ_{SPP}.

② 表面等离子激元传播距离 δ_{SPP}.

表面等离子激元的传播距离 δ_{SPP}决定于表面等离子激元波矢的虚部 k_{SPP}''，

$$k_{SPP}''=k_0\frac{\varepsilon_m''}{2(\varepsilon_m')^2}\sqrt{\left(\frac{\varepsilon_m'\varepsilon_d}{\varepsilon_m'+\varepsilon_d}\right)^3}. \tag{8.1.21}$$

表面等离子激元传播距离定义为当电磁模的功率与强度之比降到初始值的 1/e 时，表面等离子激元沿表面所通过的距离，即

$$\delta_{SPP}=\frac{1}{2k_{spp}''}=\lambda_0\frac{(\varepsilon_m')^2}{2\pi\varepsilon_m''}\sqrt{\left(\frac{\varepsilon_m'+\varepsilon_d}{\varepsilon_m'\varepsilon_d}\right)^3}. \tag{8.1.22}$$

当金属的损耗很低时，则有$|\varepsilon_m'|\gg|\varepsilon_d|$，$\delta_{SPP}$可以近似地表示为

$$\delta_{SPP}\approx\lambda_0\frac{(\varepsilon_m')^2}{2\pi\varepsilon_m''}. \tag{8.1.23}$$

图 8.9(a)和(b)所示的曲线分别是在可见光和近红外光频率范围内归一化的表面等离子激元波长及表面等离子激元传播长度与光波长的关系.采用银的介电常数,可以估算出在可见光和近红外光波长范围内银表面激发的表面等离子激元波长分别为 0.5 μm 和 1 μm,而其在空气/银界面上传播长度分别为 20 μm 和 500 μm.

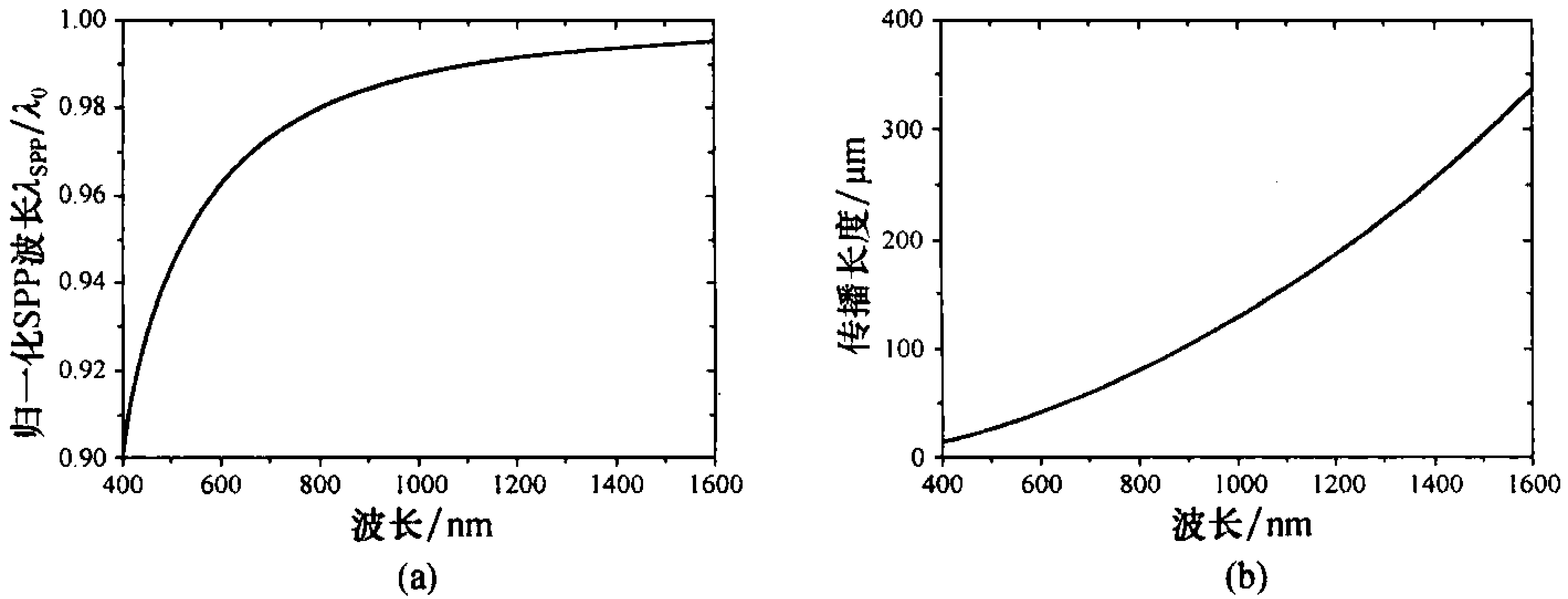


图 8.9 在可见光和近红外光频率范围内归一化的表面等离子激元波长(a)和表面等离子激元传播长度与光波长的关系(b)

那么,要使 δ_{SPP}增长,则要求金属的介电常数具有一个大的负实部 ε_m'和小的虚部 ε_m'',即低损耗的金属材料.δ_{SPP}决定了表面等离子激元元件和回路的最大尺寸上限.要增长 δ_{SPP}的主要办法是采用耦合的表面等离子激元电磁模.另外,要求表面等离子激元传播距离 $\delta_{SPP} \gg \lambda_{SPP}$,这意味着在金属表面可以通过刻蚀周期性光栅结构或各种褶皱结构(周期与 λ_{SPP}同量级)来操控表面等离子激元的传播.此时表面等离子激元电磁模容许与许多周期内精细结构发生耦合作用,散射表面等离子激元电磁模进入自由传播的光场中去.

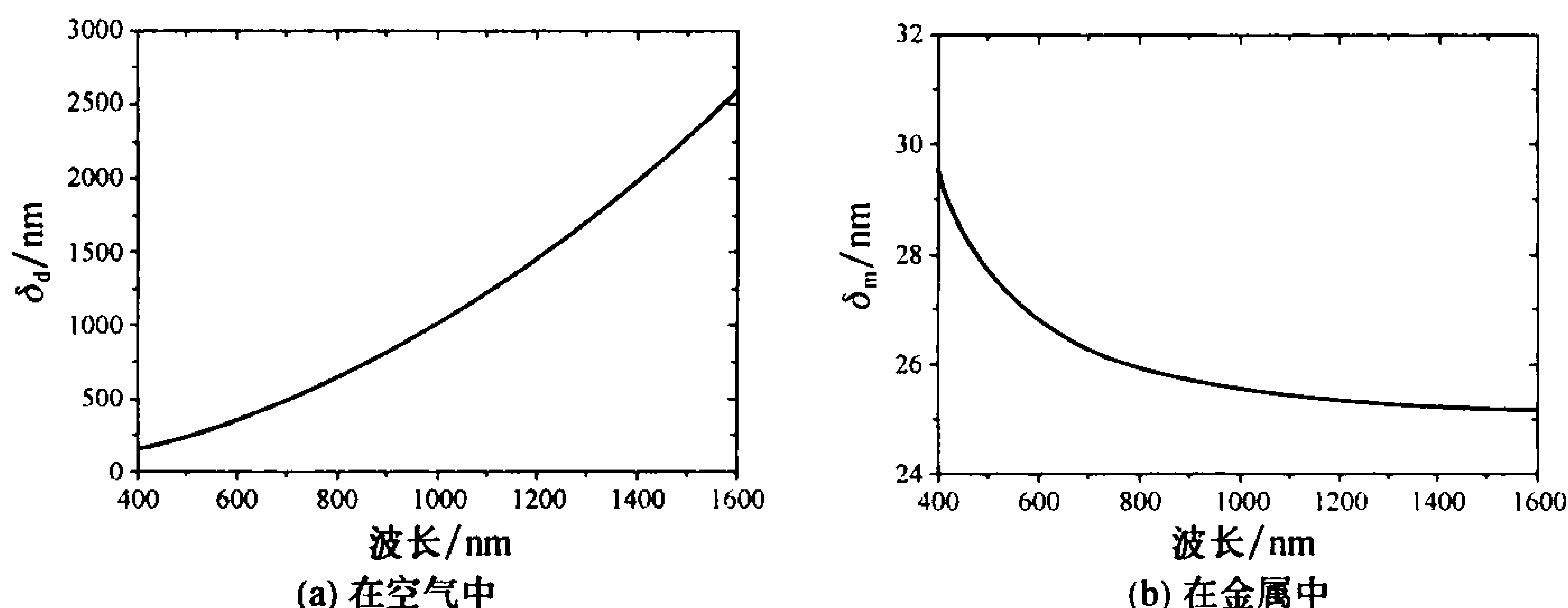
③ 表面等离子激元穿透趋肤深度 δ_d 和 δ_m.

考虑由一平坦平面构成的金属-电介质界面,利用前面推算出色散关系式(8.1.15),可以容易地计算出表面等离子激元场穿透入电介质中的趋肤深度 δ_d 和场穿透入金属中的趋肤深度 δ_m 分别为

$$\delta_d = \frac{1}{k_0}\left|\frac{\varepsilon_m' + \varepsilon_d}{\varepsilon_d^2}\right|^{1/2}, \tag{8.1.24}$$

$$\delta_m = \frac{1}{k_0}\left|\frac{\varepsilon_m' + \varepsilon_d}{(\varepsilon_m')^2}\right|^{1/2}. \tag{8.1.25}$$

图 8.10(a)和(b)所示的曲线分别表示在可见光和近红外光波长范围内表面等离子激元在空气中穿透趋肤深度 δ_d 和在金属中穿透趋肤深度 δ_m 与光波长的关系.



(a) 在空气中 (b) 在金属中

图 8.10 表面等离子激元穿透趋肤深度与光波长的关系

§8.2 表面等离子激元局域增强特性

由于表面等离子激元具有表面局域和近场增强特性，且其与光的相互作用时所呈现出的局域增强特性会随着金属表面的亚波长纳米结构的改变而发生变化，因而备受人们广泛关注. 目前国际上实现纳米尺度的光学控制主要方法之一就是基于表面等离子激元. 利用表面等离子激元可以将光学控制的维度降为二维，实现纳米尺度超衍射极限的光传输的有效调控，同时可在纳米尺度上电磁能量局域汇聚放大，实现其空间尺度受限、场空间局域增强的特点，在微纳传感探测、纳米光子器件设计及其集成等纳米光子学领域具有重要应用.

基于表面等离子激元器件的广泛研究都与其具有的近场增强局域特性密切相关. 譬如，表面增强传感，提高二极管发光效率及太阳能电池效率、纳米天线等方面的应用都是由于其空间局域近场增强效应. 由于表面等离子激元的迷人特性，再加上一些基本物理机理问题还需要进一步探索，这使得近年来表面等离子激元研究既具有很高的理论研究价值，又具有新颖应用前景.

8.2.1 表面等离子激元近场增强特性

沿金属表面传播的表面等离子激元是电磁波与电荷密度集体振荡波的混合. 它在垂直于金属表面的方向电场强度是呈指数衰减的，对应于表面等离子激元局域特性. 它的另一个独特的性质是近场光的增强. 近场光的增强程度取决于金属的介电常数、表面粗糙程度引起辐射损耗以及金属薄膜的厚度等因素.

在金属表面传播的表面等离子激元电场可以表示为[9]

$$\boldsymbol{E}_{\mathrm{SPP}} = \boldsymbol{E}_0 \exp(k_{\mathrm{SPP}} x - k_z |z|). \tag{8.2.1}$$

该式明确表明表面等离子激元表面局域特性，其电场强度沿着垂直金属表

面方向指数衰减.根据麦克斯韦方程和边界条件,可以解出理想光滑平面上表面等离子激元最大可能的场增强表示为[9]

$$\frac{\boldsymbol{E}_{\mathrm{SPP}}^{z=0}}{\boldsymbol{E}_{\mathrm{in}}}=\frac{2}{\varepsilon_{\mathrm{d}}}\frac{|\mathrm{Re}\varepsilon_{\mathrm{m}}|^{2}}{\mathrm{Im}\varepsilon_{\mathrm{m}}}\frac{a}{1+|\mathrm{Re}\varepsilon_{\mathrm{m}}|},\tag{8.2.2}$$

其中,$a=\sqrt{|\mathrm{Re}\varepsilon_{\mathrm{m}}|(\varepsilon_{\mathrm{s}}-1)-\varepsilon_{\mathrm{s}}}$,$\boldsymbol{E}_{\mathrm{in}}$是入射光的电场,$\varepsilon_{\mathrm{s}}$ 是金属膜下面用于激发表面等离子全反射棱镜的介电常数,ε_{d} 是金属膜上层介质的介电常数.由式(8.2.2)可估算出表面等离子激元场增强,对于 60 nm 厚的银膜在红光照射下增强可达 2 个数量级.表面等离子激元近场增强特性可使其适合用于高灵敏度的传感器.

8.2.2 局域表面等离子激元

表面等离子激元的另一个重要特点是局域性,即在金属/电介质界面上存在比较大的场强分布.利用表面等离子激元局域特性可以用于调控光与物质相互作用并增强材料的光学非线性,例如金属微纳结构可以用来增强表面拉曼显微镜(surface-enhanced Raman spectroscopy, SERS)信号,从而用于单分子探测[13].如果金属膜的表面非常粗糙或形成金属曲面结构(如球体、柱体等),其表面等离子激元不能以波的形式沿界面传播,而是被局域在这些结构的表面附近,这就是表面等离子激元局域化.在这种特殊的情况下,表面等离子激元也被称作局域表面等离子激元(localized surface plasmon,LSP).

如图 8.11 所示,当尺寸接近或小于光波长的金属颗粒被光照射后,其振荡电场使金属颗粒的电子云相对于核心发生位移,由于电子云和核心间库仑引力作用产生恢复力,引起电子云在核心周围的振荡,这种电子的集体振荡被称为局域表面等离子激元共振[14,15].振荡的频率主要由金属的电子密度、有效电子质量、颗粒的尺寸、形状、周围介质等因素决定.局域表面等离子激元共振在金属纳米颗粒光学性质中扮演着关键的角色.

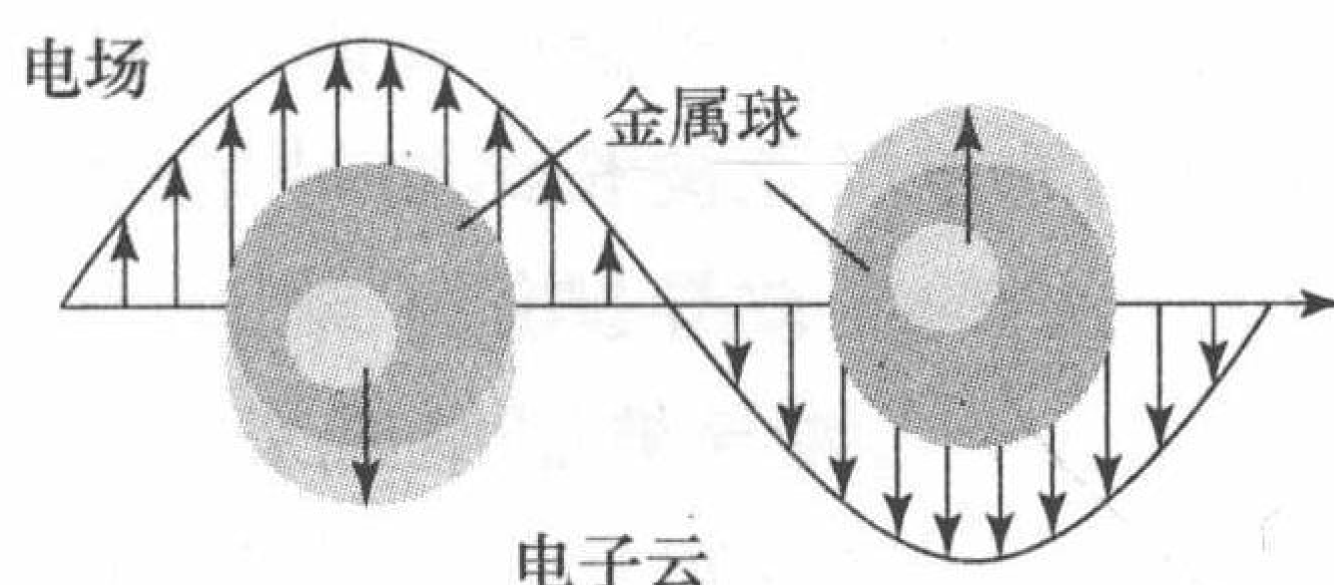


图 8.11 金属球体的表面等离子激元振荡示意图

对于金属纳米颗粒,利用局域表面等离子激元局限在一个很小区域的电场增强效应,可以使得许多光学过程的效率得到显著的提高.金属纳米材料在表面拉曼散射增强技术上的重要应用就是一个典型例证[16].现在利用这项技术,甚

至能够实现单分子信号的检测.金属颗粒的局域表面等离子激元的共振频率及其电场对金属颗粒的形状、结构和组分的明显依赖性,这可以使得其在纳米尺度上对电场进行“修饰”[17].

需要强调的是,虽然表面等离子激元和局域表面等离子激元都具有表面局域特性,但是各自色散关系的不同决定了它们是两种完全不同的激发态.表面等离子激元色散关系由式(8.1.19)决定,由电场表达式(8.1.11)可知,它是一种TM波传播模式(在平行于金属界面方向).而局域表面等离子激元是局域在各种不同形貌的金属曲面上,它的色散关系一般和其所局域的电介质形貌密切相关.

§8.3 表面等离子体激元器件

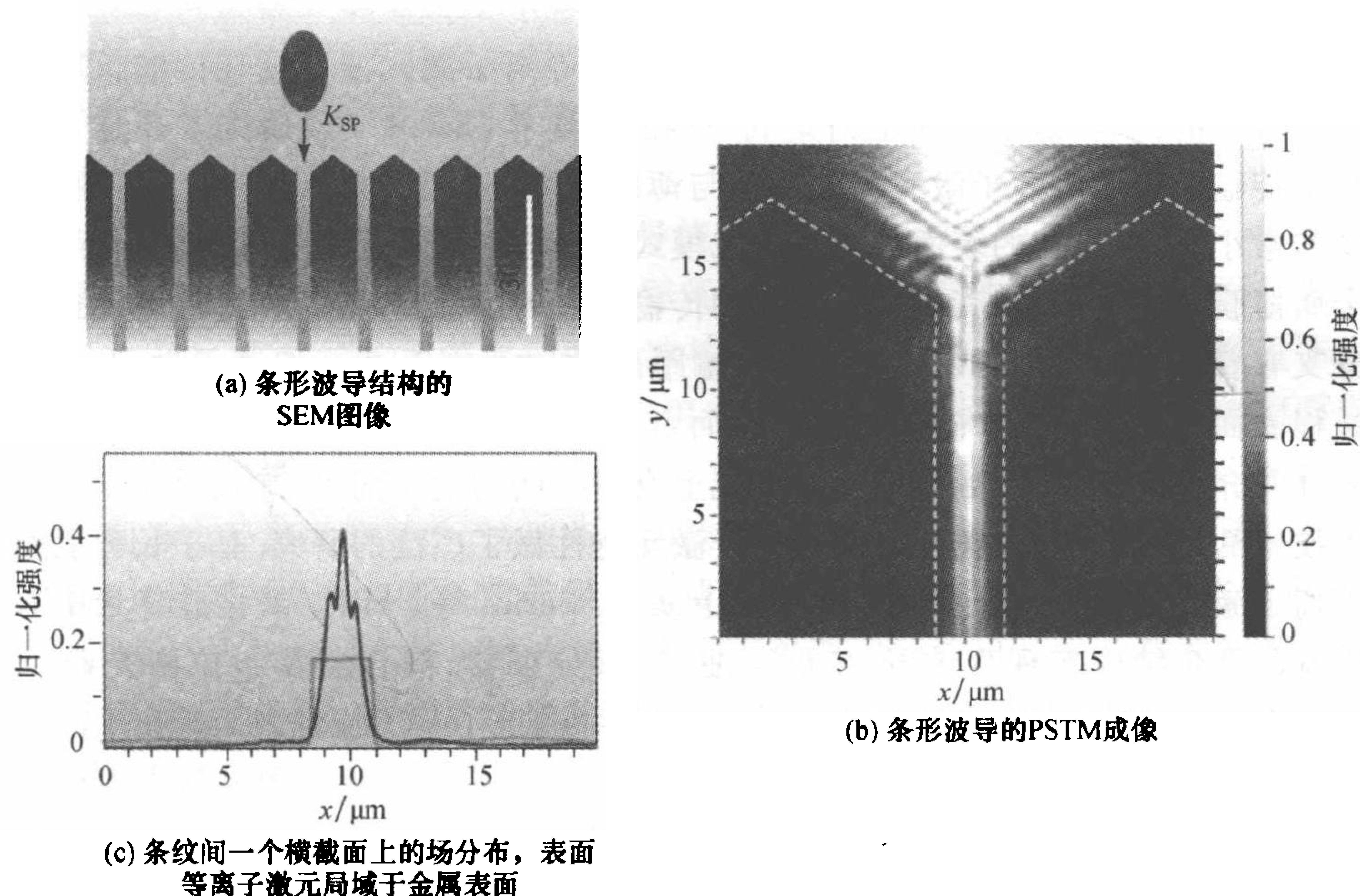
自从Ebbesen在1998年发现表面等离子激元引起超强光学透射现象[18]并在2002年发现亚波长孔径凹槽结构具有光集束功能后[19,20],相关领域涌起了对表面等离子激元器件研究的热潮.科学研究表明表面等离子激元器件能突破传统光学器件受衍射极限限制这一瓶颈.如今,人们已经认识到表面等离子激元器件已经在现代信息光电子、绿色能源、微波和太赫兹波等领域显示出重要的应用前景.常见的表面等离子激元器件包括有源(如表面等离子激元源、表面等离子激元太阳能电池和表面等离子激元增强型光发射二极管等)和无源(如表面等离子激元波导等)两类器件.

8.3.1 表面等离子激元波导

在集成光学领域,表面等离子激元的特点在于其能够突破衍射极限.表面等离子激元波导是一种能够将光信号转化为表面等离子激元,利用表面等离子激元作为信号载体并在光衍射极限以下的区域内进行传输的新型器件.要利用表面等离子激元作为波导,可以通过一定的金属亚波长结构来实现.通过解带边界条件的麦克斯韦方程,周期性的金属亚波长结构对光的汇聚和引导作用可以被模拟出来,因而为设计这种类型的光波导提供了帮助.在很多有条件的实验室,已经有很多可行的表面等离子激元波导被制作出来,表面等离子激元波导的实现和应用具有重要的意义.长远来说,它是实现全光回路(optical circuit)的基础.由于表面等离子激元可以突破衍射极限,利用表面等离子激元波导结构制作全光回路芯片成为目前研究方向.

在设计表面等离子激元波导时,人们总希望在表面等离子激元传播方向的垂直截面上的波导结构对光场的束缚性要好,以降低波导的折弯损耗和提高光子芯片中回路和元件的密度.与此同时,又希望波导能保持较低的传输损耗,以

提高表面等离子激元传播距离. 但是由于对表面等离子激元的束缚一般是通过降低模式在介质材料中空间分布来实现,这又将导致被金属所吸收的能量比例增加使得传输损耗也增加. 因此需要优化波导的几何形态,来平衡相互制约着的传输损耗和模式尺度. 图 8.12 是一个 40 nm 厚的金膜做成的表面等离子激元条形波导结构[6]. 其中图 8.12(a)是条形波导结构的 SEM 图像,图 8.12(b)是条形波导的光子隧道显微镜(photon scanning tunneling microscope, PSTM)成像,图 8.12(c)是条纹间一个横截面上的场分布,表面等离子激元局域于金属表面.



(a) 条形波导结构的SEM图像

(b) 条形波导的PSTM成像

(c) 条纹间一个横截面上的场分布，表面等离子激元局域于金属表面

图 8.12　表面等离子激元波导结构

首先提到的是金属纳米颗粒阵列结构[21,22]. 近年来由于使用化学方法来制备金属纳米颗粒在合成技术上的突破以及一些微纳米加工技术的发展,如电子束刻蚀、聚焦离子束刻蚀等,使得金属纳米颗粒阵列的制作及研究成为可能. 金属纳米颗粒中的表面等离子激元与块体金属材料和电介质界面上的表面等离子激元不同,在外界光的照射下,金属颗粒中的电子也会发生群体移动,从而使得金属颗粒中的电子密度发生重新排布,这样就在金属颗粒界面内外分别重新产生电场,形成表面等离子激元. 由于金属颗粒界面将自由电子束缚于金属颗粒内,所以电子的群体移动也就被局限于金属纳米颗粒中,这时产生的表面等离子激元被称为局域化表面等离子激元. 金属纳米颗粒局域化表面等离子激元的性

质跟金属种类、颗粒形状及环境介质都有很大关系.金属纳米颗粒阵列结构虽然能提供对光波在横向上的亚波长约束,但理论和实验结果都证明,这种波导结构由于金属的内在吸收而伴随着非常大的损耗.例如,Maier等[21]曾报道在可见光波段尺寸为90 nm×30 nm×30 nm,颗粒间距为50 nm一维银纳米颗粒阵列的实验测量损耗值高达3 dB/100 nm.这说明在光波段,金属颗粒阵列波导的传输长度只有亚微米量级.这么大的损耗限制了金属纳米颗粒阵列结构在光子集成中的进一步应用,并且这种金属颗粒阵列的形状使得这种波导难以与其他光波导相互集成.基于金属纳米颗粒阵列结构的波导器件也只有直波导、T型分支等器件在文献中被报道过.

与金属纳米颗粒阵列波导相比,长程表面等离子激元波导则在传输长度上有非常大的优势,易于实现平面集成.早在1981年,Sarid[23]就研究了在金属薄膜上传输的表面等离子激元色散特性与薄膜厚度的关系,并指出当金属薄膜厚度非常小时,所支持的表面等离子激元模式将分为两个分支.Sarid还指出,随着金属薄膜厚度的减小,其中一个分支的传输常数的虚部将趋向于零.这表明这个分支有着非常小的损耗及比较长的传输距离,称之为长程表面等离子激元.由于长程表面等离子激元伴随的损耗相对而言比较小,模式分布呈偶对称而容易被光纤等介质波导激发,因此在光子集成上有着广阔的应用前景.加上近年来微加工技术的发展,人们对长程表面等离子激元特性做了广泛的研究.基于长程表面等离子激元的多种波导都逐渐在实验上实现.Berini等[24]以二氧化硅背景中有限宽条形金结构实现直波导、S形弯曲、Y形分束器、耦合器及布拉格光栅[25]等.Boltasseva等[26]在光通信波段设计、制作并测试了基于长程表面等离子激元波分复用器,还在实验[27]上测得嵌入在苯并环丁烯(BCB)聚合物材料中厚度为100 nm的波导,在1550 nm波长下的传输损耗为6 dB/cm.此外,Boltasseva等还测量了波导形成半径为15 mm时的弯曲损耗为5 dB,并基于这种波导实现了多模干涉器件及定向耦合器.Nikolajsen等利用热光效应在嵌入聚合物材料的金属条上分别实现了光调制器[28]及光开关[29].可见,目前人们采用长程表面等离子激元波导在光子集成的应用方面取得了长足的进展.

虽然长程表面等离子激元波导的损耗随着金属膜层厚度的减小而减小,从而能获得较长的传输长度,但是却以牺牲了表面等离子激元局域性为代价.随着金属膜层厚度的减小,长程表面等离子激元场分量以倏逝波的形式向两边的介质中扩散至几个波长的数量级甚至更大,形成一种准平面波模式.于是,人们开始探寻其他类型的表面等离子激元波导结构,特别是在金属中夹着电介质的那种金属-电介质-金属(metal-insulator-metal, MIM)结构.

MIM结构对光的约束能力比较强,但是与长程表面等离子激元相比,这种结构伴随的损耗要大得多.于是,研究人员以保证亚波长约束为前提,以降低损

耗为目的，在MIM结构的基础上设计并研究了各种类型的表面等离子激元波导，包括不同金属的异质结构波导[30]、金属矩形槽波导[31,32]、金属V形槽波导[33~35]等.

在上述众多的表面等离子激元波导中，V形槽表面等离子激元波导因为具有相对低的损耗、亚波长约束以及能无弯曲损耗地通过90度弯曲而备受研究人员关注. 2006年，丹麦奥尔堡大学的Bozhevolnyi教授领导的研究小组[36]在《自然》杂志上报道了他们使用聚焦离子束刻蚀加工、并用近场扫描光学显微镜观察所得到的V形槽表面等离子激元波导研究结果. 2007年，该研究小组[37]又在实验上制作并测试了基于V形槽表面等离子激元波导的分/插复用器及布拉格光栅滤波器，从而实现了对表面等离子激元的滤波特性. V形槽表面等离子激元波导在实验上的实现是表面等离子激元在亚波长光子集成应用研究中的一个突破.

MIM类型的表面等离子激元波导能将光的模式约束到亚波长量级，但这也给该波导实验研究带来了新的困难，那就是如何将光有效地耦合到表面等离子激元波导中. 在对槽表面等离子激元波导的实验研究中，Bozhevolnyi教授研究小组采用的办法是在微操纵平台下使用锥形透镜单模保偏光纤，并在光学显微镜的辅助下将锥形光纤末端对准到V形槽表面等离子激元波导的输入端，将光耦合到等离子激元波导中. 这种办法的缺点需要非常精密的微操纵仪器及操纵经验，并且耗时久并且不一定能保证成功率，尤当V形槽的横向尺寸比较小时. Chen等[38]直接通过电子束曝光技术在金属缝隙波导输出及输入端通过接上最宽处宽度为450 nm的锥形结构的硅波导，从而可以直接把光由普通电介质波导耦合到表面等离子激元波导. 他们通过这种方法测得缝隙为150 nm宽且芯层介质为硅的金属缝隙波导在1550 nm波长下的损耗约为0.8 db/μm.

研究人员通过优化设计表面等离子激元波导结构可以尽可能地减少传输损耗，但是由于金属的内在吸收特性，要从根本上解决损耗问题需要在表面等离子激元波导中引入增益介质. 目前研究人员已经在增益介质补偿表面等离子激元波导损耗方面进行了尝试性研究[39]. 对于能提供亚波长约束的MIM类型波导，Maier[40]通过计算指出，在1550 nm波长下，假设芯层为半导体材料时(折射率3.4)，则当芯层宽度分别为100 nm及50 nm时，在MIM波导中实现无损耗传输所需要的半导体增益系数分别为1625 cm^{-1}及4830 cm^{-1}，这都在目前的半导体增益介质所能提供的增益能力范围之内. 由于MIM的损耗随着芯层介质折射率的减小而降低，因此，当芯层材料使用聚合物等低折射率材料时，所需要的增益更小. 除了半导体增益介质外，研究人员还采用染料增益[41]等途径来尝试增强表面等离子激元波导的传输. 总之，使用增益介质来降低甚至抵消表面等离子激元波导的损耗是一种比较有潜力的途径，但目前在这个领域还有许多研究

工作需要开展.

目前的研究是致力于实现全表面等离子激元回路,着眼于实现电光、全光、压电调制的表面等离子激元波导.大量数据要求从一个芯片的一端界面传送到另一端界面时,首先将电信号转换成表面等离子激元信号,沿着表面等离子激元波导传播,然后在终端处,再转换成电信号.然而目前的表面等离子激元波导的性能尚不足以承担这一重任.

困扰表面等离子激元波导研究进一步发展的主要问题来自于两个方面:① 光与表面等离子激元的耦合效率较低;② 空间局域性与传播距离(衰减)之间的矛盾.那么,对于一种有效的表面等离子激元耦合方法的研究将具有重要的意义.

8.3.2 表面等离子激元源

严格来说,表面等离子激元源是一种能够激发表面等离子激元的金属-电介质纳米结构,一般包括电致和光致表面等离子激元源两类器件.利用表面等离子激元来增强发光材料和器件量子效率的有源器件也是一种表面等离子激元源,它是目前材料发光领域中的一个研究热点.

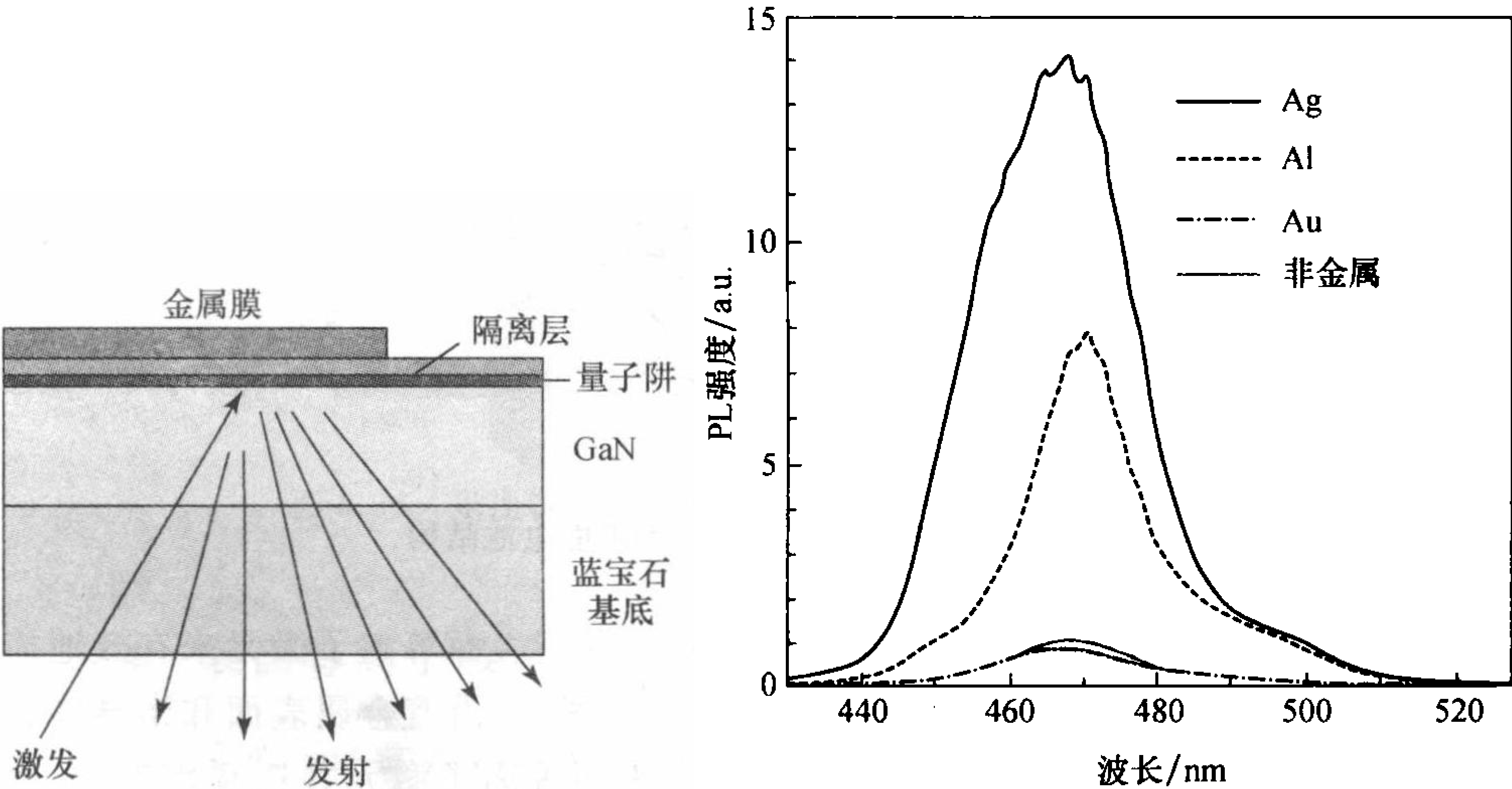
表面等离子激元一个重要特性就是它能与满足一定耦合条件的电磁波发生耦合,使电磁波能量转化为表面等离子激元能量;也能通过适当的条件,使表面等离子激元转化为普通电磁波.因此表面等离子激元可以融入一系列发光现象的研究中.

最近几年,研究者开始尝试利用金属-电介质纳米结构中的表面等离子激元高局域场特性来提高材料发光中心的光子态密度,进而提高材料的辐射跃迁几率.譬如基于 InGaN 的半导体发光二极管(LED)已成为各种固态发光装置的候选者之一,但是它面临着发光效率低的困境.利用表面等离子激元将有助于解决这一核心问题.当 InGaN/GaN 量子阱被一纳米金属银或铝薄膜覆盖时,表面等离子激元的激发将增大态密度和增强自发辐射,从而大大提高其内量子效率.T. Atay 等在实验中已经观测到激发表面等离子激元后其发光效率增强了 32 倍[42].表面等离子激元也能够用于改善激光器性能.例如,将一个金属纳米孔径加工到一个 GaAs 垂直腔表面发射激光器的顶部,可作为亚波长光学近场探针.光的近场强度和信号电压都明显地增大.

2000 年,加州理工学院研究小组分别利用理论和实验分析了表面等离子激元增强 LED 的发光特性,他们的结构包括 InGaAs/GaAs 量子阱和上下两个金属薄膜,其中上层金属薄膜通过刻蚀法制备了纳米光栅结构,理论估计该结构可以提高光抽出效率 37%,并且柏塞尔(Purcell)因子达 4.5[43].在实验中测得的刻蚀金属光栅结构的样品比没有光栅结构的样品的荧光强度提高了 46 倍,他们

将这个增强归因于这种金属微腔结构俘获了更多的泵浦光，使激发强度和效率大大提高. 2002 年，A. Neogi 等利用时间分辨荧光谱测量到了 InGaN/GaN 量子阱的复合几率，发现在与 Ag 薄膜的表面等离子激元共振耦合后提高 92 倍，并且利用费米黄金定则计算了该耦合过程，同时提出减小量子阱与金属层的距离将得到更大复合几率[44].

2004 年，加州理工学院的 A. Scherer 研究小组首次利用金属表面等离子激元成功地增强了 InGaN 量子阱结构的发光强度，并且分析了不同金属材料以及结构对耦合增强发光的影响，该实验结果发表在《自然 · 材料》(*Nature Materials*)上[45]. 图 8.13(a)，(b)所示的是其样品的结构、荧光测量方式和荧光光谱图. 为了避免距离金属膜太近而造成能量转移淬灭，量子阱与金属膜之间加入 10 nm 厚的 GaN 隔离层，由于银的表面等离子体能量与 InGaN 量子阱的发光能量最接近，所以镀银膜的样品荧光强度增加最多，其内量子效率提高了 6.8 倍，同时他们还比较了不同结构的银薄膜对荧光增强的影响，发现与薄膜结构相比，金属光栅结构耦合发光效率更高. 虽然该研究仅是初步的成果，其得到的结构仍然是光泵浦荧光的增强，并没有实现电泵增强，但它展示了将 LED 中的表面等离子激元损耗变为有益的光发射的可能性. 之后陆续出现了很多 InGaN 量子阱与表面等离子激元耦合发光的报道，比较详细地研究了表面等离子激元耦合发光的温度效应[46]、金属光栅耦合发光的极化特性[47]和金属薄膜精细结构对耦合发光的影响[48]等.



(a) 样品结构和PL光谱测量的激励/发射结构　(b) 分别涂以银、铝、金的InGaN /GaN 量子阱的PL光谱

图 8.13　PL 光谱测量

8.3.3 表面等离子激元太阳能电池

表面等离子激元一个突出特点就是局部增强效应,这种局部增强效应已经应用于提高薄膜太阳能电池效率.薄膜太阳能电池能够大幅度降低传统太阳能的成本,但是在薄膜晶体硅太阳能电池中,为了增强光的吸收,提高电池效率,光在薄膜中的束缚至关重要.一种实现薄膜太阳能电池光捕获的新方法是采用能够支持表面等离子激元的金属-电介质纳米结构.通过对这些纳米结构做适当的处理、光能够被聚集和"折叠"进入薄的半导体层,从而增加光吸收.

表面等离子激元太阳能电池是一种利用表面等离子激元来捕获太阳光从而极大地提高太阳能电池转换效率的新型器件.例如,Catchpole 等[49]在 2008 年报道了他们利用金属纳米颗粒的局域化表面等离子激元在增强硅太阳能电池吸收效率的工作.他们发现,表面等离子激元能够在几乎整个太阳光光谱上增强薄膜太阳能电池的频谱响应.

人们感兴趣的是利用表面等离子激元散射和耦合特性来提高单节平面薄膜太阳能电池的光吸收效率,许多新型的表面等离子激元太阳能电池结构设计考虑到增加光的局域和散射的金属纳米结构.图 8.14 所示为一种新型的表面等离子激元太阳能电池结构,它是一种采用了一个非常薄的量子点层来加强光吸收的表面等离子激元太阳能电池[50].量子点被嵌入在 MIM 表面等离子激元波导中,入射光进入表面等离子激元波导后,会在银和量子点层界面激发传播表面等离子激元.在表面等离子激元耦合作用下,从而提高了光吸收效果.

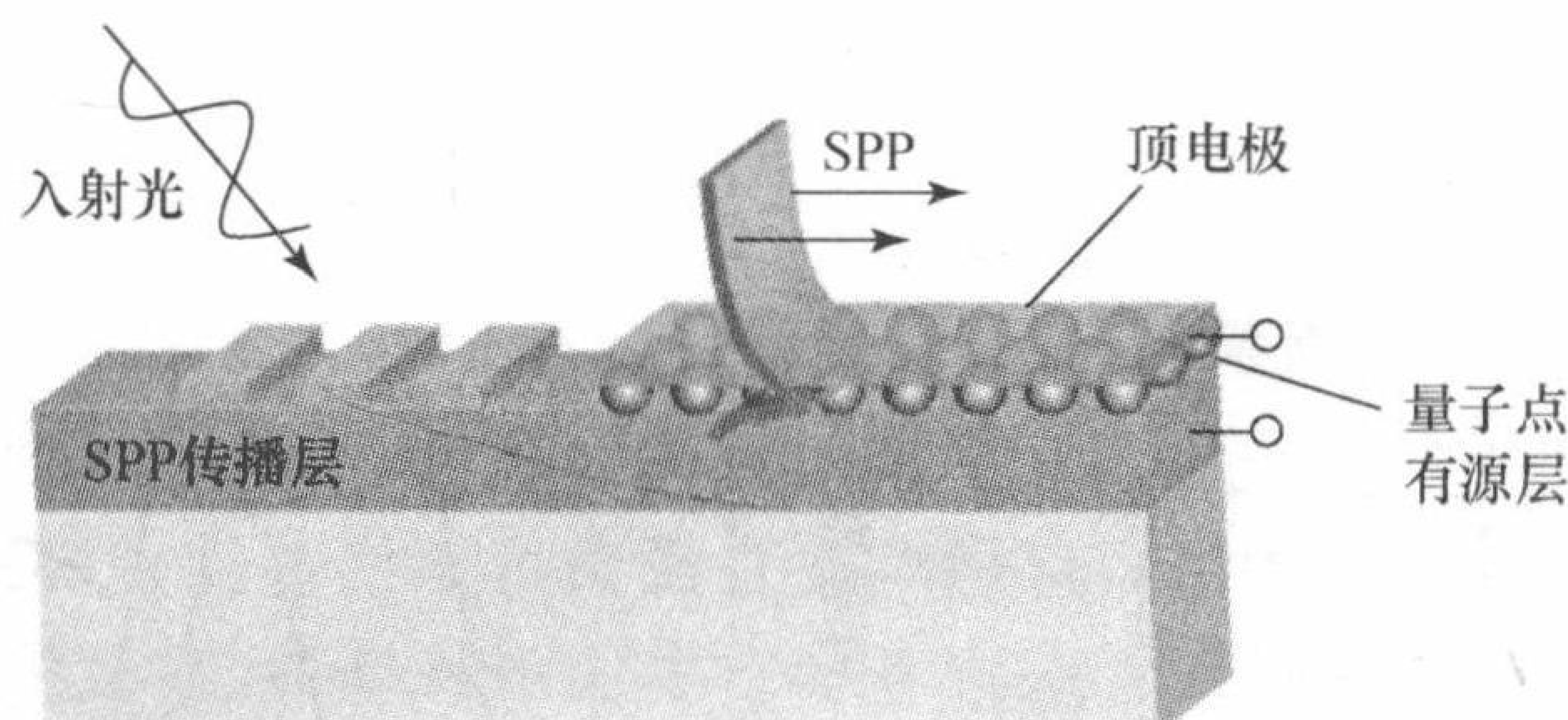
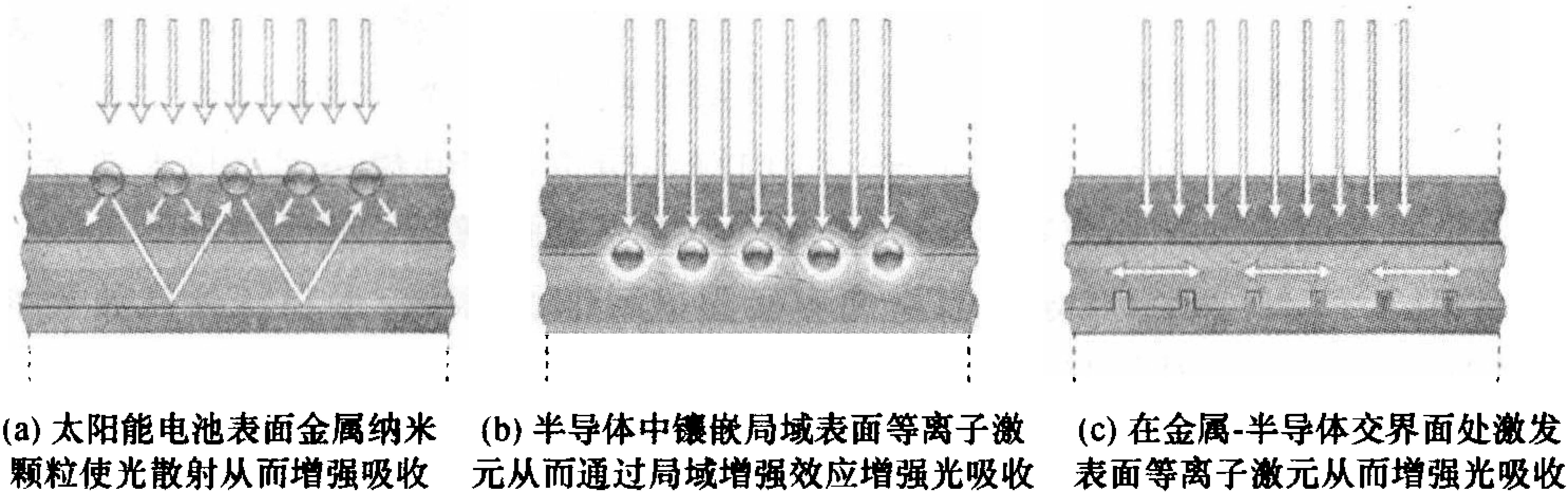


图 8.14 表面等离子激元太阳能电池结构

表面等离子激元太阳能电池的主要特点是利用表面等离子激元来有效地捕获太阳光.其工作机制是将光转换成表面等离子激元(沿着金属表面和半导体吸收层传播的特殊电磁波),如图 8.15 所示.在表面等离子激元共振频率附近,表面等离子激元在垂直方向上衰减非常明显,表面等离子激元场被束缚在远小于波长的界面上.在这种结构中,入射光能流方向变了 90°,光沿着太阳能电池的横截面传播,这就导致了光程有几个数量级的提高.不过由于存在着入射光波和

表面等离子激元动量的不匹配，于是在金属-半导体吸收层界面必须设计特殊的结构以达到这种匹配.比较可行的就是将背表面设计成光栅结构.除此之外，由于存在表面等离子激元共振频率，于是设计耦合结构使之更少依赖于频率，同时还有偏振结构，入射角度等成为一系列挑战.



(a) 太阳能电池表面金属纳米颗粒使光散射从而增强吸收　(b) 半导体中镶嵌局域表面等离子激元从而通过局域增强效应增强光吸收　(c) 在金属-半导体交界面处激发表面等离子激元从而增强光吸收

图 8.15　薄膜太阳能电池中增强光吸收的表面等离子激元结构

图 8.16 所示的是 S. Pillai 等[51]采用银纳米岛结构分别在 SOI 和硅晶片基底构成的表面等离子激元硅太阳能电池.该研究结果表明采用银纳米粒子激发的局域表面等离子激元是一种能够提高硅薄膜和晶片为基底的太阳能电池效率的方式之一，它避免了由于硅表面发生变形时而导致电子空穴复合增加的问题.由于金属纳米粒子的存在，采用这种方法可以把硅膜的厚度降低至 1.5 μm 且还具有更好的捕集光的效果.

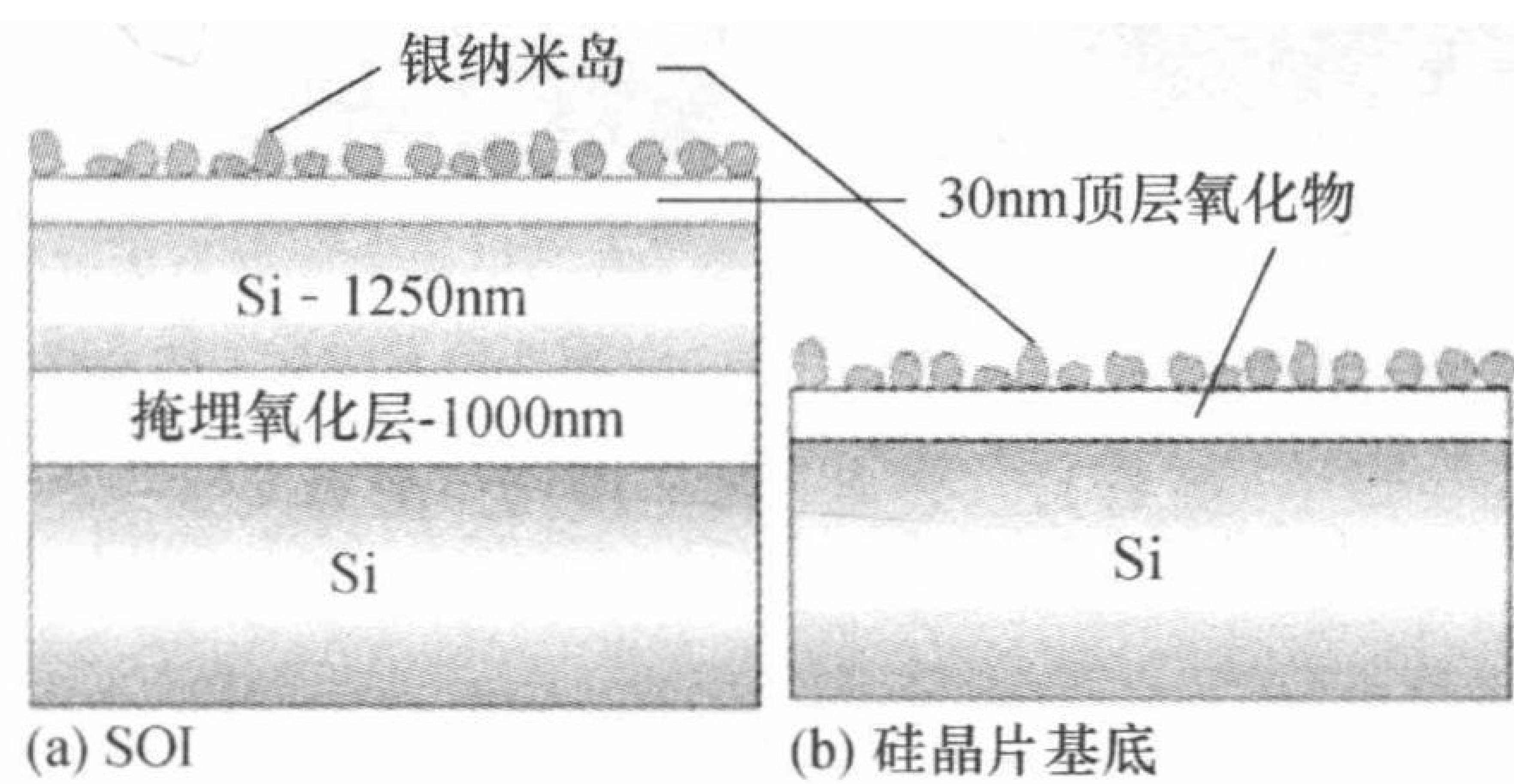


图 8.16　采用银纳米岛表面等离子激元硅太阳能电池结构

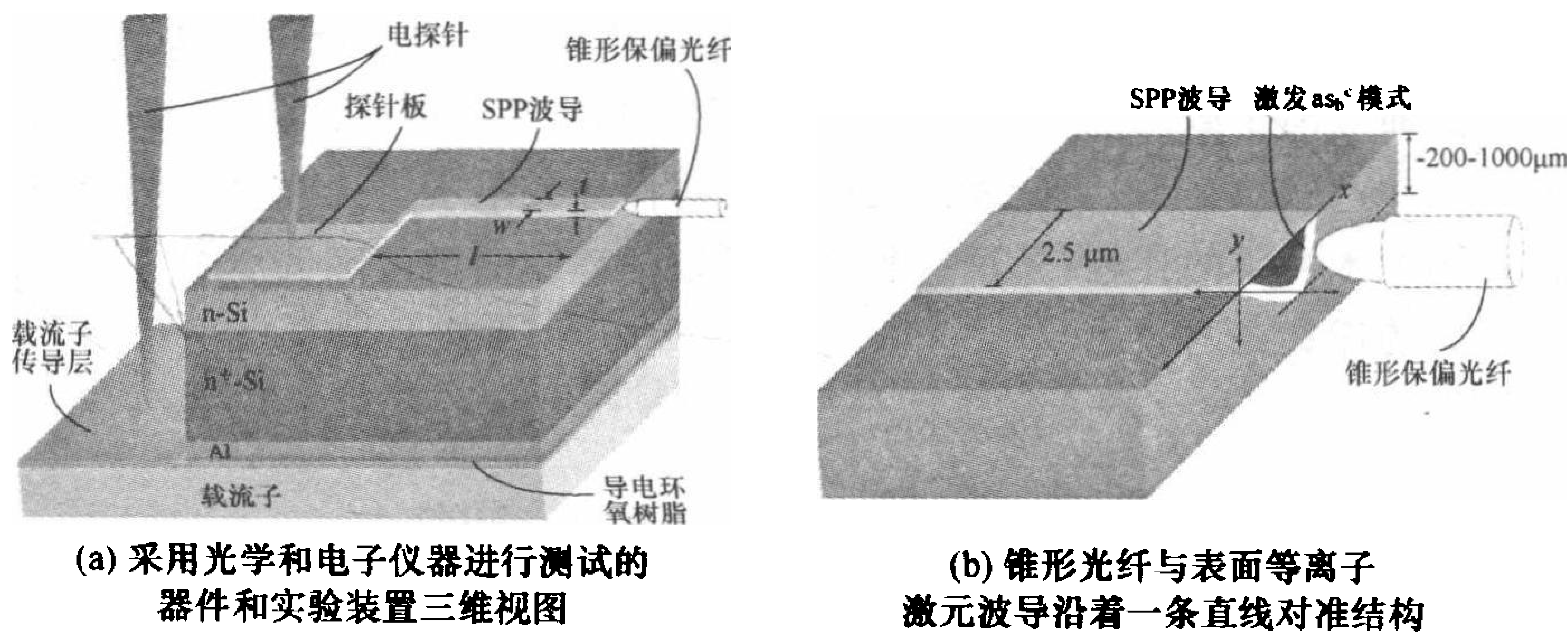
8.3.4 表面等离子激元探测器

表面等离子体激元具有极高的表面和体的敏感特性.这些特性导致其在光探测器等方面具有广泛的应用，目前表面等离子激元探测器的研究成为一个热点.表面等离子体激元探测器是一种引入了能够激发表面等离子激元的金属-电介质纳米结构，利用表面等离子激元局域增强效应来优化普通的光电探测器的

光吸收效率.其工作原理与普通的光电探测器的工作原理本质上是一样的,都是基于光电效应.如在亚波长孔径周围加工叉指形式的金属双栅,增强了光透射、吸收和响应.

譬如,人们通常为平板金属选择一轻掺杂半导体作电介质覆层在两种材料界面处形成肖特基接触,这种结构会赋予更多的光电子学功能特性.肖特基二极管结构作为一个红外探测器之所以被广泛研究是因为其光子能量低于半导体的带隙.至今为止,人们还在理论上探讨采用对称和非对称硅覆层把支持表面等离子激元的金属条波导和肖特基势垒探测器集成为表面等离子激元探测器.

2010 年,A. Akbari 等[52]首次在本实验中制作了一种应用在红外光波长内的表面等离子激元肖特基探测器,如图 8.17 所示.该器件主要是由能够支持表面等离子激元且强烈局限在金属-半导体界面中的硅和金属条形成的肖特基接触构成.低于硅带隙的红外光波段光辐射检测的工作机理是通过其内部的光电效应来实现.表面等离子激元探测器的研制开辟了一种新的集成技术,有望在短距离光互连方面具有广泛的应用.



(a) 采用光学和电子仪器进行测试的器件和实验装置三维视图

(b) 锥形光纤与表面等离子激元波导沿着一条直线对准结构

图 8.17 表面等离子激元探测器

§8.4 总 结

随着理论研究的深入和现代微加工技术的进步,对支持表面等离子激元金属微结构体系的研究已形成了一门新兴学科方向,即表面等离子体光子学.由于表面等离子激元具有独特的光学特性,在新型光源、纳米光学成像、纳米光刻、太阳能电池、生物传感器以及负折射材料等方面有着重要的应用前景,成为当前国内外学者重视的热点研究领域.

表面等离子激元是一种特殊电磁表面波模式,它沿着两种介质(满足特定的条件)的分界面传播,而在垂直于界面方向则以指数形式向两侧的介质内衰减.

一般来说，表面等离子激元具有以下三个特点：

① 场分布在垂直界面方向是高度局域的，是倏逝波，在金属场分布比在电介质中分布更集中，一般分布深度与波长量级相同；

② 在平行于表面的方向，场是可以传播的，由于损耗存在，在传播过程中会有衰减，传播距离有限；

③ 表面等离子激元色散曲线通常处在电磁波色散曲线的右侧，在相同频率的情况下，其波矢大于自由空间电磁波的波矢.

同时，表面等离子激元光子学的前景和挑战为硅基表面等离子激元器件提供了难得的新机遇. 具体包括：

① 硅基表面等离子体激元光子芯片. 基于表面等离子激元的发展有望研发硅基表面等离子激元光子芯片，用作超低损耗的光子互连元件. 硅基表面等离子激元光子芯片具有输出、输入端口，这些端通过表面等离子激元波导、耦合器，可实现光学方法连接到传统的衍射受限的光子器件上.

② 硅基表面等离子激元调制器和开关元件. 要实现表面等离子激元调制器及开关功能，则需要对表面等离子激元进行主动调控. 如，通过改变温度、磁场、电场或者借助于外部光的激发，进行开关操作.

③ 硅基表面等离子激元源.

总的说来，表面等离子激元研究是一个很有趣的工作，目前有许多值得研究的方向，有许多激动人心的结果，除以上介绍的应用外，在其他很多领域都有着重要的应用. 随着微纳米加工技术的发展，将会有越来越多的硅基表面等离子激元器件进入市场，服务人类.

参考文献

[1] Wood R W. On a remarkable case of uneven distribution of light in a diffraction grating spectrum. Proc. Phys. Soc. London, 1902, 18: 269.

[2] Fano U. The theory of anomalous diffraction gratings and of quasi-stationary waves on metallic surfaces (Sommerfeld's waves). Opt. Am., 1941, 31: 213.

[3] Ritchie R H. Plasma losses by fast electrons in thin films. Phys. Rev., 1957, 106 (5): 874.

[4] Kretschmann E, Raether H. Radiative decay of non-radiative surface plasmons excited by light. Z. Naturforschung, 1968, 23A: 2135.

[5] 何赛灵，戴道锌. 微纳光子集成. 北京：科学出版社，2010.

[6] Barnes W L, Dereux A, Ebbesen T W. Surface plasmon subwavelength optics. Nature, 2003, 424: 824.

[7] Raether H. Surface Plasmons. Berlin: Springer, 1998.

[8] Thio T, Lezec H F, Wolff P A, et al. Surface-plasmon-enhanced transmission through

hole arrays in Cr films. J. Opt. Soc. Am. B, 1999,16: 1743.

[9] Zayats A V, et al. Nanooptics of surface plasmon polaritons. Phys. Rep., 2005, 408 (324): 131.

[10] Reather H. Surface Plasmons on Smooth and Rough Surfaces and on Gratings. Berlin: Springer-Verlag, 1988.

[11] Barnes W L. Surface plasmon-polariton length scales: a route to sub-wavelength optics. J. Opt. A: Pure Appl. Opt., 2006, 8: S87.

[12] 顾本源. 表面等离子体亚波长光学原理和新颖效应. 物理, 2007, 36(4): 280.

[13] Otto A, Mrozek I, Grabhorn H. Surface-enhanced Raman scattering. Phys. Condens. Mat., 1992, 4: 1143.

[14] Kreibig U, Vollmer M. Optical Properties of Metal Clusters. Berlin, Heidelberg: Springer-Verlag, 1995.

[15] Kelly K L, Coronado E, Zhao L L, et al. The optical properties of metal nanoparticle: the influence of size, shape and dielectric environment. J. Phys. Chem., 2003, 107: 668.

[16] Jin P, Tazawa M, Xu G. Reversible tuningofsurfaceplasmon resonance of silver naoparticles using a thermochromic matrix. J. App. Phys., 2006, 99: 096106.

[17] Xu G, Tazawa M, Jin P, et al. Wavelength tuning of surface plasmon resonance using dielectric layers on silver island films. Appl. Phys. Lett., 2003, 82: 3811.

[18] Ebbesen T W, Lezec H J, Ghaemi H F, et al. Extraordinary optical transmission through sub-wavelength hole arrays. Nature, 1998, 391: 667.

[19] Grupp D E, Lezec H J, Thio T, et al. Beyond the bethe limit: tunable enhanced light transmission through a single sub-wavelength aperture. Adv. Mat., 1999, 11 (10): 860.

[20] Thio T, Pellerin K M, Ebbesen T W, et al. Enhanced light transmission through a single subwavelength aperture. Opt. Lett., 2001, 26 (24): 1972.

[21] Quinten M, Leitner A, Krenn J R, et al. Electromagnetic energy transport via Linear Chains of Silver Nanoparticles. Opt. Lett., 1998, 23(17): 1331.

[22] Maier A, Kik P G, Atwater H A, et al. Local detection of electromagnetic energy transport below the diffraction limit in metal nanoparticle plasmon waveguides. Nat. Mat., 2003, 2: 229.

[23] Sarid D. Long-range surface-plasma waves on very thin metal films. Phys. Rev. Lett., 1981, 47 (26): 1927.

[24] Charbonneau R, Lahoud N, Mattiussi G, et al. Demonstration of integrated optics elements based on long-ranging surface plasmon polaritons. Opt. Exp., 2005, 13 (3): 977.

[25] Charbonneau S, Charbonneau R, Lahoud N, et al. Demonstration of bragg gratings based on long ranging surface plasmon polariton waveguides. Opt. Exp., 2005, 13

(12): 4674.

[26] Boltasseva A, Bozhevolnyi S, Sondergaar I, et al. Compact Z-add-drop wavelength filters for long-range surface plasmon polaritons. Opt. Exp., 2005,13 (11): 4237.

[27] Boltassev A, Nikolajsen T, Leosson K, et al. Integrated optical components utilizing long-range surface plasmon polaritons. J. Lightwave Technology, 2005,23 (1): 413.

[28] Nikolajsen T, Leosson K, Bozhevolnyi S. In-line extinction modulator based on long-range surface plasmon polaritons. Opt. Commu., 2005, 244: 455.

[29] Nikolajsen T, Leosson K, Bozhevolnyi S. Surface plasmon polariton based modulators and switches operating at telecom wavelengths. Appl. Phys. Lett., 2005,85 (24): 5833.

[30] Wang B, Wang G P. Simulations of nanoscale interferometer and array focusing by metal heterowaveguides. Opt. Exp., 2005, 13 (26): 10558.

[31] Pile D F P, Ogawa T, Gramotnev D, et al. Two-dimensionally localized modes of a nanoscale gap plasmon waveguide. Appl. Phys. Lett., 2005,87(26): 1114.

[32] Veronis G, Fan S. Guided subwavelength plasmonic mode supported by a slot in a thin metal film. Opt. Lett., 2005,30 (24): 3359.

[33] Bozhevolnyi S I, Volkov V S, Devaux E, et al. Channel plasmon-polariton guiding by subwavelength metal grooves. Phys. Rev. Lett., 2005, 95: 046802.

[34] Moreno E, Garcia-Vidal F J, Rodrigo S G, et al. Channel plasmon-polaritons: modal shape, dispersion, and losses. Opt. Lett., 2006,31 (23): 3447.

[35] Pile D F P, Gramotnev D K. Plasmonic subwavelength waveguides: next to zero losses at sharp bends. Opt. Lett., 2005,30(10): 1185.

[36] Bozhevolnyi S, Volkov V S, Devaux E, et al. Channel plasmon subwavelength waveguide components including interferometers and ring Resonators. Nature, 2006,440: 508.

[37] Volkov V S, Bozhevolnyi S I, Devaux E, et al. Wavelength selective nanophotonic components utilizing channel plasmon polaritons. Nano. Lett., 2007,7(4): 880.

[38] Chen L, Shakya J, Lipson M. Subwavelength confinement in an integrated metal slot waveguide on silicon. Opt. Lett., 2006,31(14): 2133.

[39] Nezhad M, Tetz K, Fainman Y. Gain assisted propagation of surface plasmon polaritons on planar metallic waveguides. Opt. Exp., 2004,12 (17): 4072.

[40] Maier S. Gain-assisted propagation of electromagnetic energy in subwavelength surface plasmon polariton gap waveguides. Opt. Commu., 2006,258: 295.

[41] Noginov M A, Zhu G, Bahour M, et al. Enhancement of surface plasmons in an Ag aggregate by optical gain in a dielectric medium. Opt. Lett., 2006, 31 (20): 3022.

[42] Atay T, Song H, Nurmikko A V. Strongly interacting plasmon nanoparticle pairs: from dipole-dipole interaction to conductively coupled regime. Nano. Lett., 2004,4: 1627.

[43] Vuckovic J, Loncar M, Scherer A. Surface plasmon enhanced light-emitting diode. IEEE J. Quant. Electr. 2000,36 (10): 1131.

[44] Neogi A, Lee C W, Everitt H O. Enhancement of spontaneous recombination rate in a quantum well by resonant surface plasmon coupling. Phys. Rev. B,2002,66: 153305.

[45] Okamoto K, Niki I, Shvrtser A, et al. Surface-plasmon-enhanced light emitters based on InGaN quantum wells. Nature Materials, 2004. 3: 601.

[46] Lu Y C, Chen C Y, Yeh D M, et al. Temperature dependence of the surface plasmon coupling with an InGaN /GaN quantum well. Appl. Phys. Lett. , 2007, 90(19): 3103.

[47] Shen K C. Polarization dependent coupling of surface plasmon on a one-dimensional Ag grating with an InGaN/GaN dual-quantum-wells structure. Appl. Phys. Lett. , 2008. 92(1): 3108.

[48] Chen C Y, Yeh D M, Lu Y C, et al. Dependence of resonant coupling between surface plasmons and an InGaN quantum well on metallic structure. Appl. Phys. Lett. , 2006, 89(20): 3113.

[49] Catchpole K R, Polman A. Plasmonic solar cells. Opt. Express,2008,16: 21793.

[50] Aatwater H A, Polman A. Plasmonics for improved photovoltaic devices. Nature Materials,2010,9: 205.

[51] Pillai S, Catchpole K R, Trupke T, et al. Surface plasmon enhanced silicon solar cells. J. Appl. Phys. , 2007,101(9): 3105.

[52] Akbari A, Tait R N, Berini P. Surface plasmon waveguide Schottky detector. Opt. Exp. ,2010,18 (8): 8505.

第九章 硅基光电子器件工艺与系统集成

随着加工工艺技术的进步和对更高集成度的需求,人们开始高度关注器件制备技术的发展.对于光电子集成器件而言,工艺误差将会导致器件性能偏移理论设计值,工艺水平将直接影响器件的性能.如纳米光波导的制备通常需要采用电子束曝光或深紫外光刻技术.电子束曝光过程中的邻近效应和紫外光刻过程中的衍射效应等会导致制备出来的波导宽度偏离设计值,从而导致器件性能偏移.另外,刻蚀造成的侧壁粗糙度是器件损耗的主要来源,波导的侧壁粗糙度将直接影响器件的通光性能.因此,研究和探索高精度的微纳加工技术非常重要.

这一章我们主要介绍硅基光电子器件的制备工艺,特别是其中关键的薄膜制备技术、光刻/电子束光刻技术和刻蚀技术.

§9.1 平面工艺

9.1.1 平面工艺概述

平面工艺是20世纪60年代发展起来的一种非常重要的半导体加工技术.该工艺是在硅半导体芯片上通过氧化、光刻、离子注入、刻蚀等一系列流程,制作出器件.由于器件都是在芯片表面一层附近处,整个芯片基本上保持是平坦的,故称相应的制作工艺为平面工艺.现在平面工艺已经是制造各种半导体器件与集成电路的基本工艺技术.

光电子器件的制备通常也是采用与平面工艺兼容的加工技术.标准的微加工工艺流程如图9.1所示,其中最关键的工艺步骤是薄膜生长沉积、光刻和刻蚀.我们将在下面各节中分别对这几个关键技术进行详细的介绍.

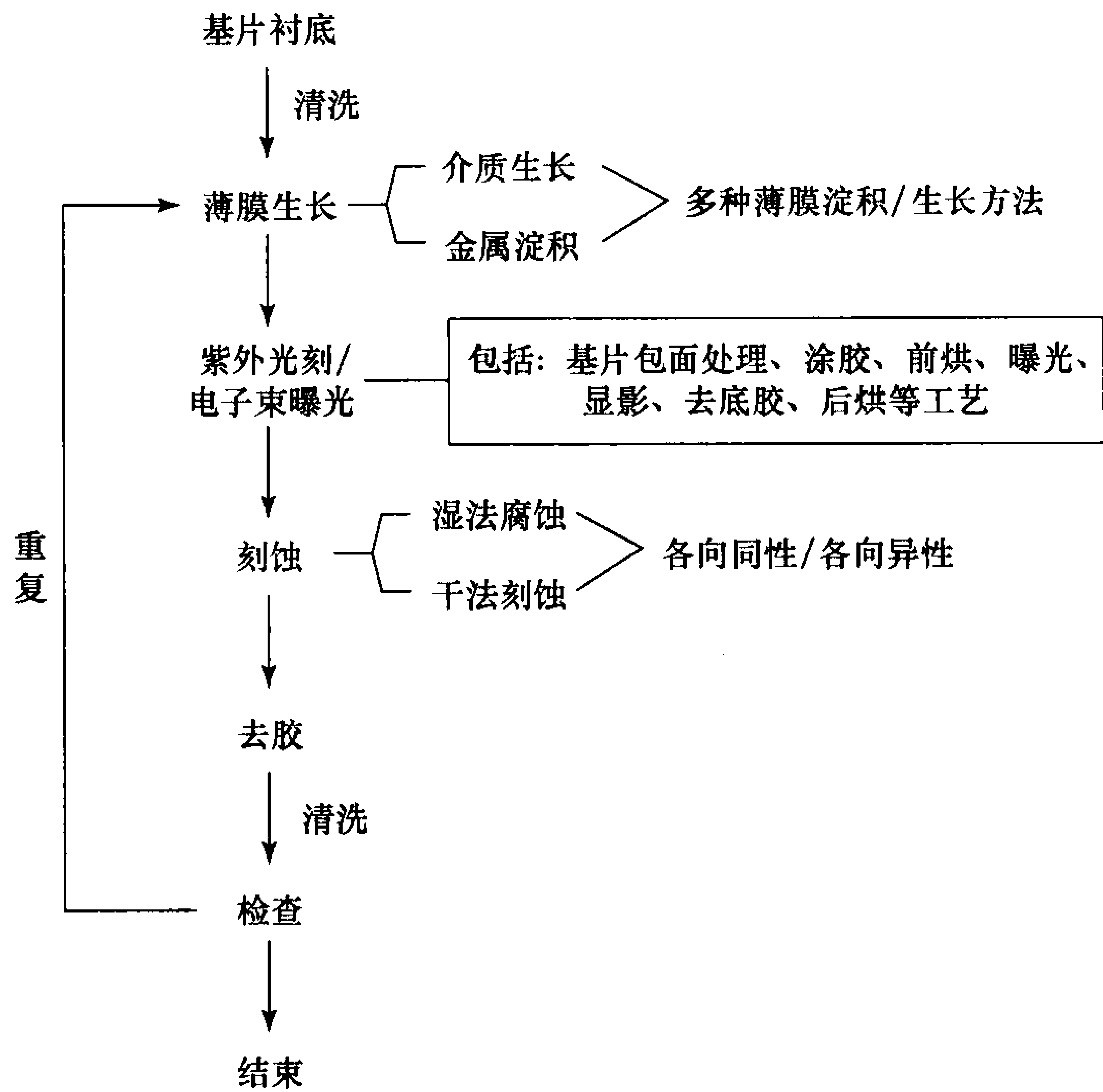


图 9.1 标准的微加工工艺流程

9.1.2 典型光波导器件工艺流程

这里，我们以一个最简单的例子：SOI 直波导的制备，来说明微纳光子器件制备基本的工艺流程. 有些结构还需要用到薄膜生长、金属剥离(lift-off)、掺杂、退火等工艺，这里我们不一一讨论. 典型微纳光子器件制备工艺流程如表 9.1 所示.

表 9.1 典型微纳光子器件工艺流程

序号	工艺流程	工艺步骤	所用设备	说明	结果
1		硅片清洗	通风橱	一般用丙酮、IPA、去离子水(电学器件有时需用到强酸去除金属离子)	
2		烘烤	热板或烘箱	除去硅片表面水分	

（续表）

序号	工艺流程	工艺步骤	所用设备	说明	结果
3	HMDS	涂增粘剂	涂胶机或热板、烘箱等	有些胶与硅片表面粘附性不好，需要涂增粘剂 HMDS	
4		涂胶	涂胶机	不同旋胶速度得到不同胶厚	
5		前烘	热板或烘箱	去除胶膜中残存的溶剂，温度时间比较重要	
6		（深）紫外曝光	曝光机（如 MA-6 等）	曝光光刻胶表面指定区域. 速度快，但精度较低	
		电子束曝光	电子束曝光设备	曝光电子束光刻胶表面指定区域. 精度高，但速度较慢	
7		显影	通风橱	显影得到光刻图形	
8		图形检查	显微镜、SEM、AFM 等	检查图形，判断光刻参数是否合理	

（续表）

序号	工艺流程	工艺步骤	所用设备	说明	结果
9		后烘	热板或烘箱	也叫坚膜，增加光刻胶抗刻蚀性	
10	~RF	刻蚀	RIE 或 ICP	转移图形从光刻胶到硅片	
11		去胶	通风贮	去除剩余光刻胶，对难去除的胶还要用到等离子体去胶(descum)	
12		图形检查	显微镜、SEM、AFM 等	检查器件结构，与设计值比较	

注：1. 序号带下划线的步骤为可选步骤；

2. 此表以正胶为例，负胶将得到相反图形；

3. 若用"剥离"工艺，将得到相反图形.

§9.2 薄膜制备

9.2.1 薄膜制备技术

总体上来说，薄膜的制备方法分为物理方法和化学方法两大类.

1. 物理方法

制备薄膜所用的物理方法通常叫做物理气相沉积(physical vapour deposition，PVD)方法，根据沉积过程的差异又分为蒸发和溅射两种方法.

热蒸发气相沉积法又叫真空蒸镀方法，它是在真空下加热蒸发材料(蒸镀材料)，使其蒸发粒子沉积在基板表面形成薄膜的一种方法. 按照加热方式，热蒸发气相沉积方法分为电阻加热、闪电加热、激光加热和电子束加热等真空蒸镀方

法.电阻加热方法是用钨、钼等灯丝直接加热蒸发材料,或者把蒸发材料放在坩埚里间接加热.高熔点材料一般用电子束加热方法.石墨等物质用电弧沉积方法.

溅射沉积(sputtering deposition)方法是用离子束轰击靶物质,与靶表面原子或分子发生弹性或非弹性碰撞.结果部分靶表面原子成分子蒸发出来,沉积在基板上形成薄膜.溅射装置也分几种类型:离子束溅射装置在离子室中形成离子,然后进入到溅射室溅射靶物质;其他溅射装置都采用辉光放电阴极溅射方法,它们没有单独的离子室.溅射沉积方法直接通过溅射蒸发靶物质形成薄膜,因此高熔点和合金等用溅射方法很容易实现薄膜化.

2. 化学方法

化学气相沉积(CVD)是半导体工业中应用最为广泛的用来沉积多种材料的技术,包括大范围的绝缘材料,大多数金属材料和金属合金材料.从理论上来说,它是很简单的:两种或两种以上的气态原材料导入到一个反应室内,然后他们相互之间发生化学反应,形成一种新的材料,沉积到晶片表面上.

CVD技术常常通过反应类型或者压力来分类,包括低压CVD(LPCVD)、常压CVD(APCVD)、亚常压CVD(SACVD)、超高真空CVD(UHCVD)、等离子体增强CVD(PECVD)、高密度等离子体CVD(HDPCVD)以及快热CVD(RTCVD),还有金属有机物CVD(MOCVD)等.

薄膜沉积最重要的两个参数是沉积速率和台阶覆盖.沉积速率与沉积材料、沉积温度、腔内气压以及基底与靶的位置关系等因素有关.沉积速率通常用石英晶体速率指示器测量,位于腔体内部.该器件是一个谐振器板,它可以在谐振频率下振荡,工作时测量其振荡频率.因为晶体顶部有材料的沉积,所外加的质量将会使得振荡频率发生偏移,通过测量该频率偏移便可得出沉积速率.但沉积足够厚的材料后,谐振器失效,不再出现尖锐的谐振,此时需要更换速率指示器.台阶覆盖表示沉积过程中薄膜对台阶的覆盖能力和均匀性.台阶覆盖与沉积方法、设备、沉积温度、气压等因素有关.由于沉积的方向性,有时也可通过旋转基底的方法改善台阶覆盖的均匀性.

9.2.2 物理气相沉积

1. 蒸发

常用的有三种蒸发系统:电阻加热、电感系统和电子束系统.

电阻加热采用最简单的加热源,一个带电源输入的高真空腔,加上沉积材料和可调变压器就可以构成一个简单的蒸发台.由于加热灯丝至少要加热到使材料蒸发的温度,因此电阻加热一个有关的问题是灯丝的蒸发和出气.如果要蒸发高熔点的材料,常常没有合适的电阻加热元件.

另一种方法是电感加热坩埚.蒸发材料放在一个一般由氮化硼制成的坩埚中,一个金属线圈绕在坩埚上,通过该线圈施加射频(radio frequency, RF)功率,射频在坩埚内感应出涡流电流使材料升温.线圈本身可采用水冷,保持温度低于100℃,有效地避免了线圈材料的损耗.

电感加热可用于提高坩埚温度,以蒸发难熔金属,但坩埚本身材料的玷污仍是一个重要的问题.采用只加热材料而冷却坩埚的方法可避免这种影响,实现此目的的常用方法是电子束(electron beam)蒸发系统.这是目前使用最广泛的蒸发系统之一,它包括一个加热钨丝环,围绕着一根相对钨丝处于高偏压的材料细棒周围.从钨丝射出的电子束通过磁场偏转轰击材料靶,从而使材料蒸发.大多数的电子束蒸发台是在坩埚下面的电子枪发射出有一定强度的高能束流,灯丝这样放置能减少灯丝材料向基底表面的沉积,减小由灯丝材料带来的玷污.一个强磁场将束流弯曲270°,使电子束打到材料表面.电子束可以扫描材料,以熔化大部分的材料表面,材料加热部分被有效地包裹在较冷的部分中,从而避免坩埚本身材料带来的玷污.虽然电子束蒸发可尽量避免灯丝和坩埚材料带来的污染,但同时会带来电子束辐射损伤,因此在对电子束辐射敏感的基底材料(如电子束胶)表面沉积时,应考虑到电子束辐射损伤带来的影响.典型的电子束蒸发设备如图9.2所示.

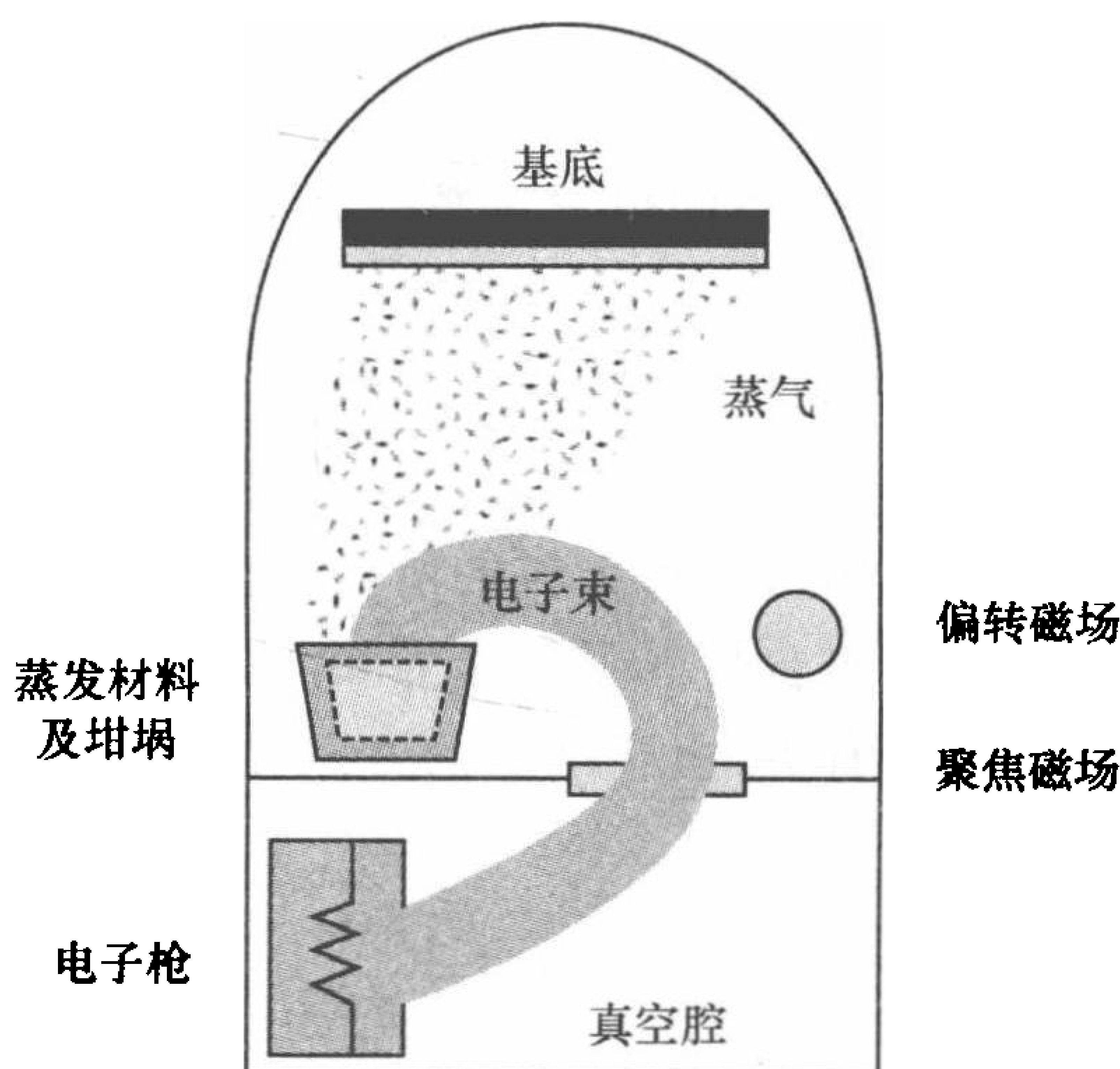


图9.2 典型电子束蒸发装置图

电子束蒸发由于腔内高真空度,因此蒸发气体的方向性很好.在对台阶进行沉积时,台阶的侧壁很难沉积到材料,这一特性使得电子束蒸发在剥离工艺中得到广泛的应用.

2. 溅射

溅射是不使用蒸发来进行材料沉积的主要替代方法，其系统简图如图 9.3 所示. 溅射的台阶覆盖比蒸发好，辐射缺陷远小于电子束蒸发，在制作复合材料膜和合金时性能更好，这些性能使得溅射技术成为许多硅基工艺的最佳选择.

溅射系统是在真空腔内产生等离子体并且使高能离子轰击溅射靶，溅射靶含有所需要沉积的材料. 溅射时，将靶材料而不是基底放置在具有最大离子流的电极上，这一点和刻蚀系统相反. 为了收集尽可能多的出射原子，一般阴极与阳极距离很近，通常小于 10 cm.

由于该工艺的物理特点，可用溅射沉积的材料种类很多. 在溅射金属材料时，由于靶材是导电的，简单的直流溅射具有较大的溅射速率，故一般用直流溅射. 而在溅射绝缘材料(如 SiO_2)时，应当使用 RF 等离子体. 另外一点值得注意的是当材料是化合物或者合金时，所沉积材料的化学匹配与靶材料可能会略有不同，但这种差别很小，通常我们可以忽略. 因为当材料溅射速率不同时，基底沉积到比份更多的高速率材料，而相反靶的面会积聚(剩余)比份更多的低速率材料. 当溅射时，由于靶表面的低速率材料密度更大，使得沉积薄膜的成分重新接近于靶体材料.

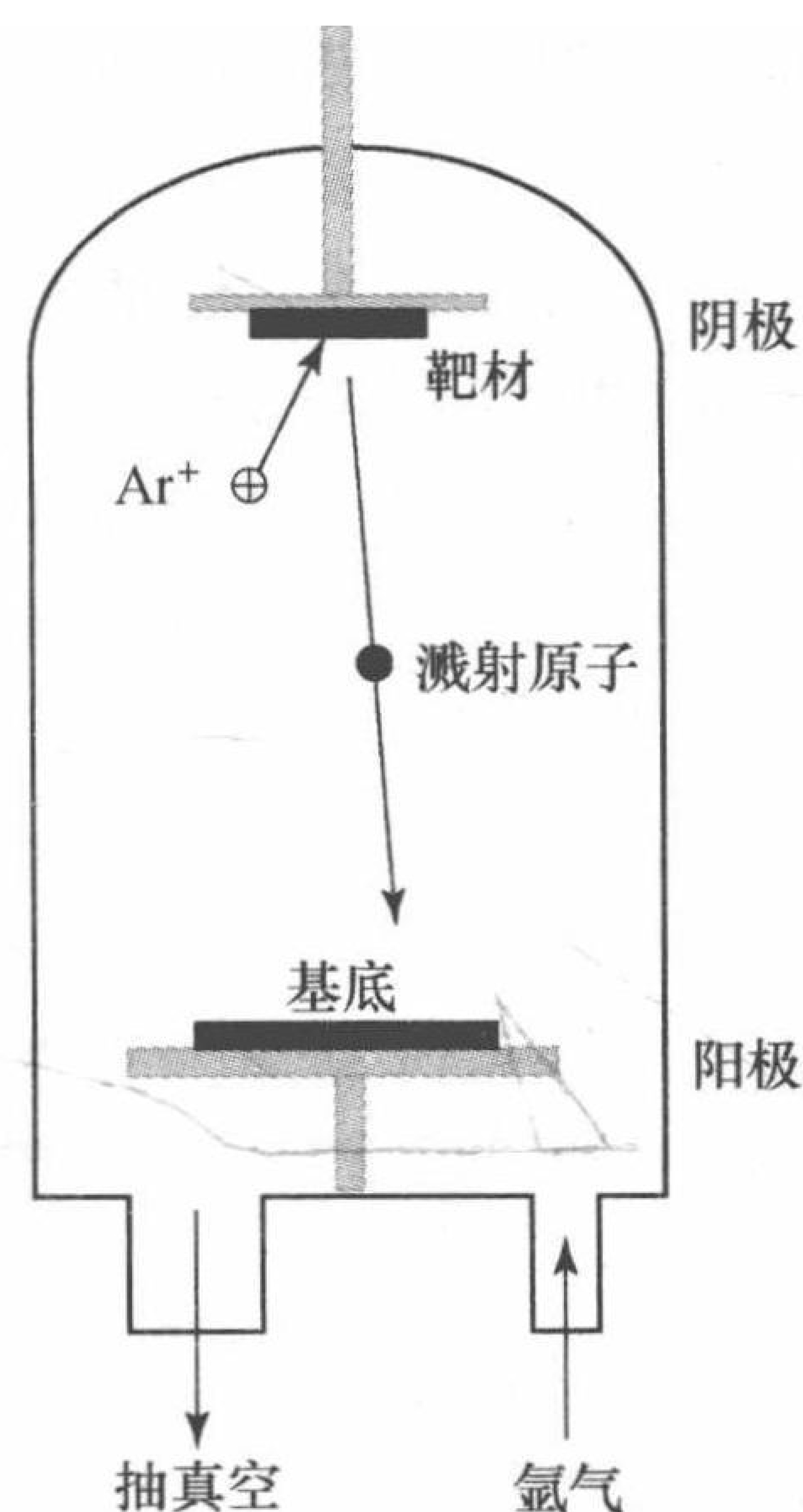


图 9.3　溅射系统简图

9.2.3　化学气相沉积

上一小节讨论了基于物理过程的薄膜沉积方法，即蒸发和溅射. 把它们称为物理气相沉积是因为这些工艺不涉及化学反应. 虽然大多数金属薄膜是使用这些方法沉积的，但他们存在着与台阶覆盖相关的一些主要问题，而且不太适合于绝缘薄膜或半导体薄膜的沉积. 这一小节主要讨论基于化学反应的薄膜沉积方法，即化学气相沉积(CVD).

一般来说，化学气相沉积的过程包含以下几步：① 反应气体从腔体入口向基底附近输运；② 气体反应生成系列次生分子；③ 反应物运输到基底表面；④ 发生化学反应，生成需要沉积的材料；⑤ 气体副产物解吸附并离开基底表面；⑥ 副产物离开反应腔体. 评价和表征化学气相沉积的指标主要有：沉积速率、薄膜均匀性、薄膜应力、台阶覆盖能力和薄膜组分等. 与前面讨论的物理沉积方法相比，化学气相沉积具有更好的台阶覆盖能力.

最早的 CVD 工艺是在大气压下进行的(即常压 CVD，APCVD)，由于反应速

率快，系统简单，特别适于介质沉积，但其缺点是均匀性差，容易形成颗粒. 为了提高沉积的均匀性，工艺设计必须设计成使反应严格受沉积动力学控制. 此时不使用稀释气体，而使用低气压(0.1～1.0 Torr)来降低气相成核，这项工艺通常称为低压 CVD(LPCVD). 一个典型的应用是利用 LPCVD 沉积多晶硅，沉积温度介于 575～650℃，反应气体为硅烷，典型多晶硅沉积速率为 10～100 nm/min. 同时，LPCVD 在沉积的过程中还可以用来进行掺杂.

在许多的应用中，需要在非常低的基底温度下沉积薄膜. 例如，在铝上沉积氮化硅. 为了适应较低的基底温度，一种最常用的方法是等离子体增强化学气相沉积(PECVD). PECVD 系统利用离子轰击表面，带来的附加好处是对次生物质提供能量，使得他们在没有高基底温度条件下进一步沿表面扩散. 因此，该工艺在填充小的几何结构方面非常好. PECVD 主要用来沉积介质材料，如二氧化硅，氮化硅等常用材料一般使用 PECVD 来沉积. 图 9.4 为典型 PECVD 系统原理图和在 Si 表面采用 PECVD 沉积 1 μm 二氧化硅后的电镜照片.

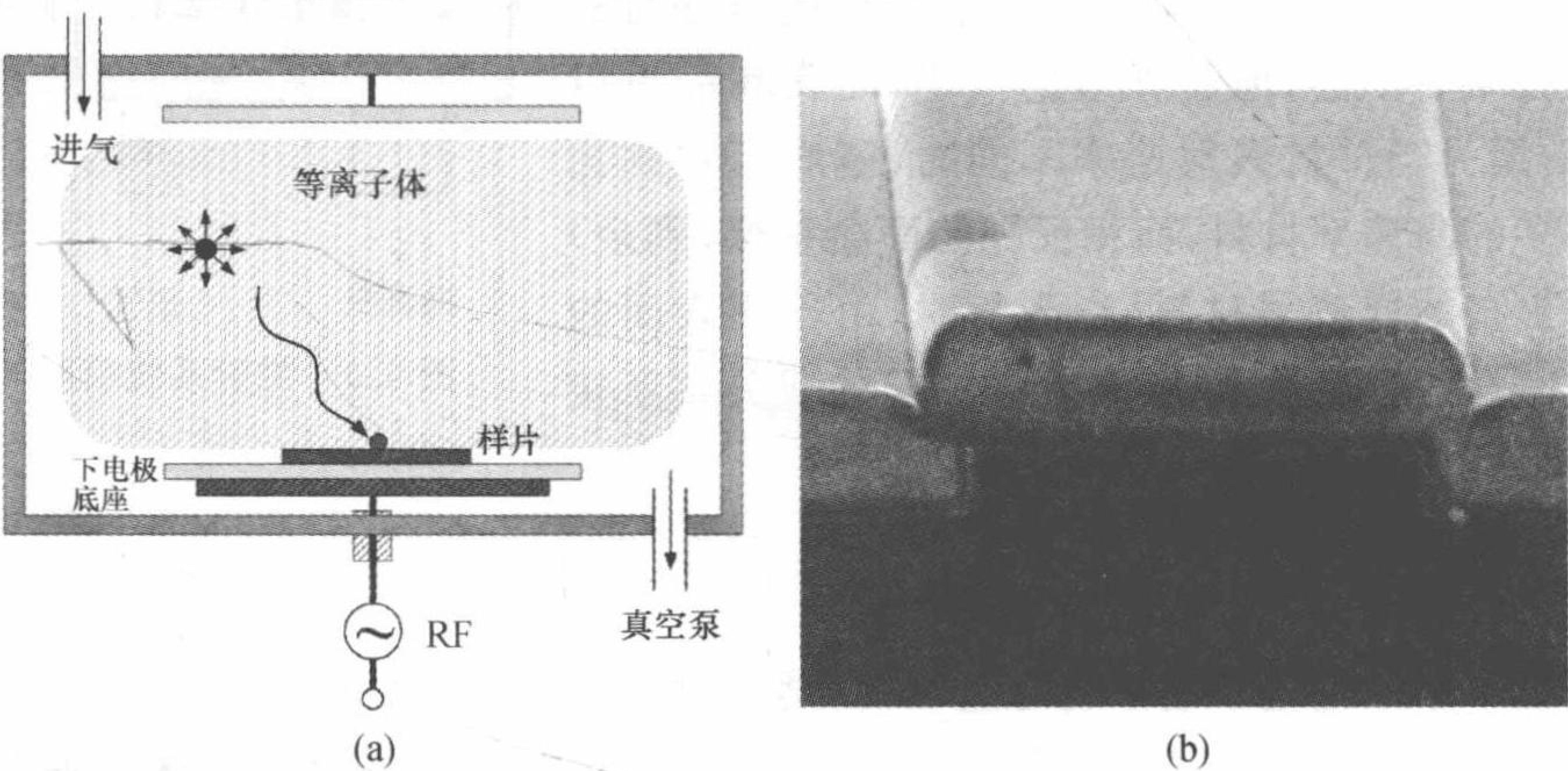


图 9.4 PECVD 系统原理图(a)；Si 表面采用 PECVD 沉积 1 μm 二氧化硅的电镜照片(b)

§9.3 光刻技术

9.3.1 光刻技术简介

光刻是指经曝光系统将预制在掩模板上的图形按所要求的位置，精确传递到预涂在晶片表面或介质层上的光刻胶薄层上的过程. 光刻是微纳结构制备基

本工艺中最关键的，也是获得图形结构的第一步．光刻确定了器件的关键尺寸，光刻过程中的错误可造成图形歪曲或对准不好，最终可转化为对器件的性能产生影响[1]．

微纳加工是制作微米和纳米尺寸量级微小结构的加工技术的总称，包括光刻技术和非光刻技术．典型的光刻技术包括光学曝光光刻(接触式光刻、接近式光学光刻和光学投影曝光光刻)、电子束光刻、X射线光刻、离子束光刻、原子力光刻、等离子体光刻等；非光刻技术主要是指机械加工、化学组装和纳米压印技术．其中光学投影曝光光刻、电子束光刻已成为大规模集成电路制造的主要手段．极紫外光刻、X射线光刻、离子束光刻、干涉光刻技术处于研究和样机阶段．纳米压印技术、原子力光刻和等离子体光刻尚在前沿研究之中．

一般的光刻流程包括基片表面处理、涂胶、前烘、曝光、显影、后烘(可选)等步骤．光刻技术工艺流程是一个较为复杂的过程．各工艺环节互相影响，互相制约，下面对每一步的作用进行简单的说明．

(1) 硅片预处理．

硅片预处理主要是对硅片进行清洗、干燥和硅片表面成膜处理表面的某些状态和特性，使光刻胶与硅片表面有良好的粘附性．

(2) 涂胶．

涂胶就是在硅片表面覆盖一层光刻胶．根据工艺要求光刻胶应达到一定的厚度，胶厚必须均匀，而且无灰尘、杂质玷污和涂胶缺陷，如针孔等．涂胶的方法很多，有喷雾涂胶、流动涂胶、浸润涂胶、滚动涂胶、旋转涂胶等，最常用的是旋转涂胺法．从涂胶厚度的一致性、稳定性来说，旋转涂胶比其他涂胶方法要好．涂胶表面的均匀性和一致性直接影响刻蚀图形的分辨力和精度．一个厚度相差很大、严重凹凸不平的胶面甚至连曝光、显影的条件也难于达到，当然就不会有好的刻蚀结果．为了保证光刻的质量，涂胶厚度必须适当．如果胶过厚，则光波通过率减小，图像的分辨力将变低，刻蚀线条的精度也将变差；胶过薄，虽然图像分辨力提高，线条精度也好，但引起的问题是随着胶厚的减小，针孔密度几乎呈指数增加．

(3) 前烘．

前烘的主要目的是消除光刻胶中的溶剂，增强胶膜与硅片之间的粘附性，保证曝光的重现性和显影时成像良好．涂好的光刻胶中，常含有很多溶剂，这将使胶膜的致密性以及胶膜与硅片表面的粘附性都较差，从而使显影液的溶解速度不稳定，同时在微加工过程中容易造成物理损伤．为此必须对涂好光刻胶的胶膜在一定温度下进行烘焙，即前烘．

(4) 曝光．

曝光是用一定波长和强度的光波透过掩模，有选择地照射光刻胶，使受照射

部分的光刻胶发生光化学反应，经显影后就可获得与掩模上相应的图形．曝光是微细加工技术中最重要的工艺环节．国内外在这方面的理论研究和新技术开发十分活跃，目的是为了适应大规模集成电路生产所要求的高分辨力、高精度、高成品率的工艺要求．

(5) 显影．

显影就是把经过曝光的硅片置于特定的显影液中，把不需要的光刻胶部分溶解去除的过程．光刻胶有正胶和负胶两种，正胶和负胶在显影后获得相反的图形．对负性光刻胶，一般采用有机溶剂做显影剂，将未曝光部分的胶溶解掉．正性光刻胶则采用无机碱或有机碱的弱碱性水溶剂做显影剂，将曝光部分的胶溶解去除．显影的方法有浸入法和喷淋法两种．喷淋法是使低压高分散的显影液流经胶表面，不仅使获得的图像更为清晰，还能减少负性胶的高分子的溶胀，利于进行高分辨图像的刻蚀，目前已为越来越多的生产部门所采用．光刻胶显影的工艺条件，由其显影特性决定．这种溶解反应速率不仅与感光性物质的浓度有关，也与显影溶液的种类、性质、显影温度、时间、显影方法以及前烘和曝光等条件有关．

(6) 后烘．

后烘又称为坚膜，就是将显影液漂洗后的硅片在一定的温度下进行烘焙．后烘的目的是为了挥发掉残存的光刻胶溶剂，进一步提高胶膜与硅片表面的黏附能力，并使胶膜致密坚固．后烘的温度和时间要适当选择．若后烘不足，溶剂没有去净，胶膜不够致密和坚固，抗蚀性差，在蚀刻时会发生浮胶或严重侧蚀等问题．若后烘过头，会使胶膜发脆，从而产生因热膨胀而翘曲、剥落及图像形变等情况，或在蚀刻时发生钻蚀或浮胶．温度过高时，光刻胶会因热分解而失去抗蚀能力．后烘的条件应根据光刻胶的种类、聚合物的结构性能以及蚀刻工艺的要求而确定．后烘一般在高温烘箱中进行，要注意烘箱内温度分布的均匀性．

如前面所讲，正胶和负胶在显影后获得相反的图形，如图 9.5 所示．当刻蚀深度较小时，可直接用光刻胶作为后续刻蚀工艺的掩模来刻蚀基底材料．在刻蚀深度较大或光刻胶作为刻蚀掩模选择比不好时，一般使用剥离技术制作金属掩模来进行刻蚀，因为对于常用半导体来说，金属是很好的等离子体刻蚀掩模材料．采用剥离工艺还要注意以下几点：① 剥离工艺获得的图形和直接刻蚀获得的图形结构是相反的；② 为了更好地剥离，显影后的光刻胶图形为上窄下宽形貌比较好，一般可采用大的曝光剂量来获得；③ 金属蒸发的厚度一般不宜超过胶厚的 1/3，否则剥离很难进行；④ 有些金属（如 Au）与基底粘附性不好，需要先蒸镀一层钛（Ti）或铬（Cr）来保证其剥离时不会脱离；⑤ 金属蒸发的方向性也会影响剥离效果，一般方向性越好越利于剥离．刻蚀和剥离是最常用的两种光刻后续工艺，我们要根据需要进行选择．

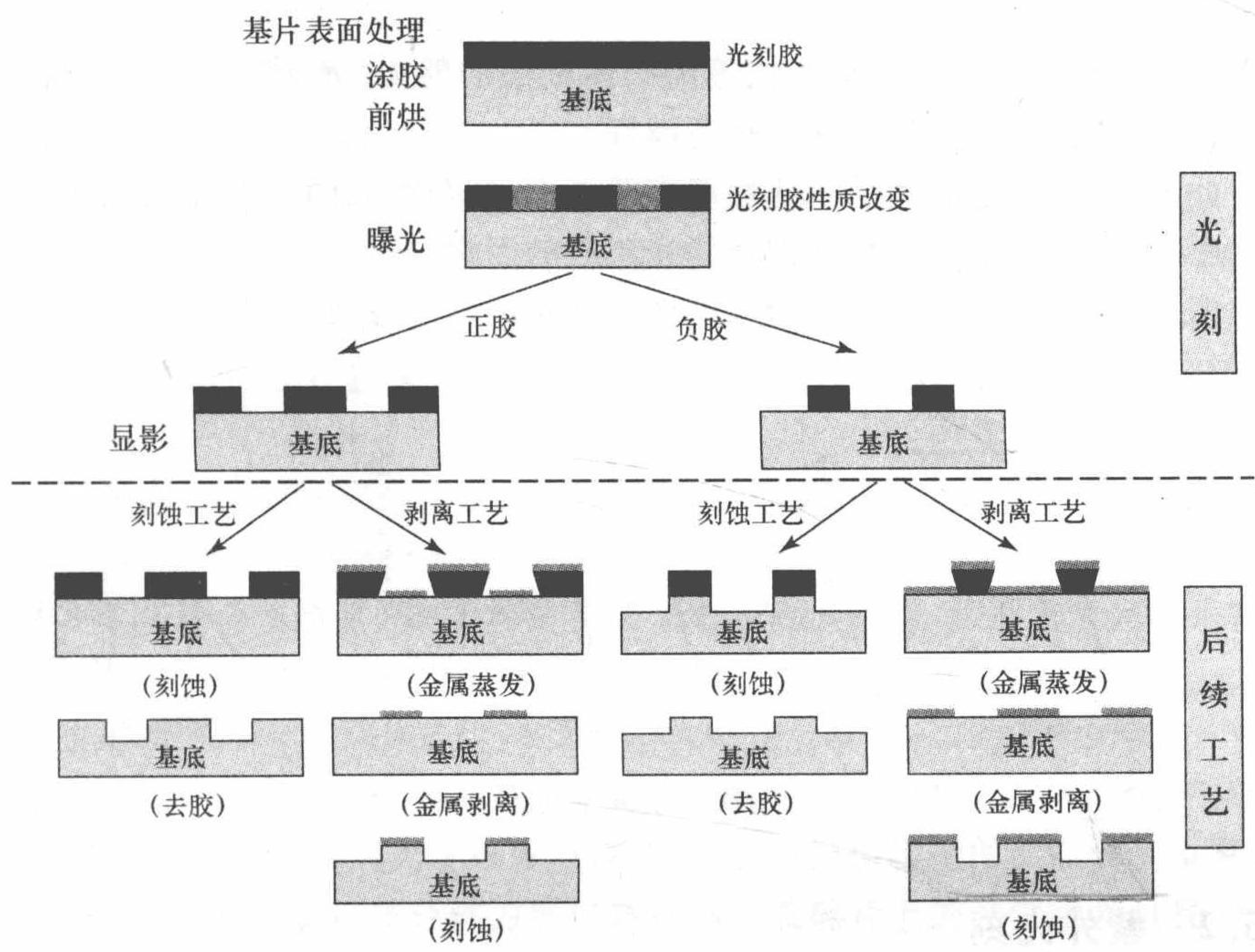


图 9.5　典型光刻及后续工艺流程

光刻胶是一种对特定波长光波敏感的有机化合物. 根据其对光波的不同响应特性可分为两类：正性光刻胶和负性光刻胶. 正性光刻胶一般在曝光下发生断链反应，变得易溶于显影液，得到和曝光图案相同的图形；而负性光刻胶一般在曝光下发生交联反应，变得难溶于显影液，得到和曝光图案相反的图形.

光刻胶有以下几个重要的指标参数：

(1) 灵敏度.

灵敏度指的是在光刻胶上产生一个良好的图形所需要的光能量. 光刻胶的灵敏度越小，所需要的曝光剂量越小，曝光速度就越快.

(2) 对比度.

对比度是表征光刻胶从曝光区域到非曝光区域过渡的陡直程度. 对比度越好，形成的图形侧壁越陡峭，所得到的图形分辨率越高.

(3) 分辨率.

光刻胶的分辨率可以说是一个综合参数. 它主要由光刻胶本身的分子特性，对比度和胶厚决定的. 通常薄胶具有更高的分辨率，但其后续工艺(如刻蚀)比较困难，因此一般要根据实际条件选择胶厚. 在满足条件的情况下尽量选择较薄的光刻胶.

(4) 抗刻蚀性.

抗刻蚀性一般是指光刻胶抗等离子体刻蚀的能力.光刻后续工艺一般是刻蚀,我们希望光刻胶的抗刻蚀性越高越好.

此外,光刻胶还有其他一些性能需要考虑,如与基底的粘附性、表面张力、保质期限等.这些也都是影响我们选择光刻胶的因素.

图 9.6 表征了光刻胶正负性、灵敏度和对比度等参数.

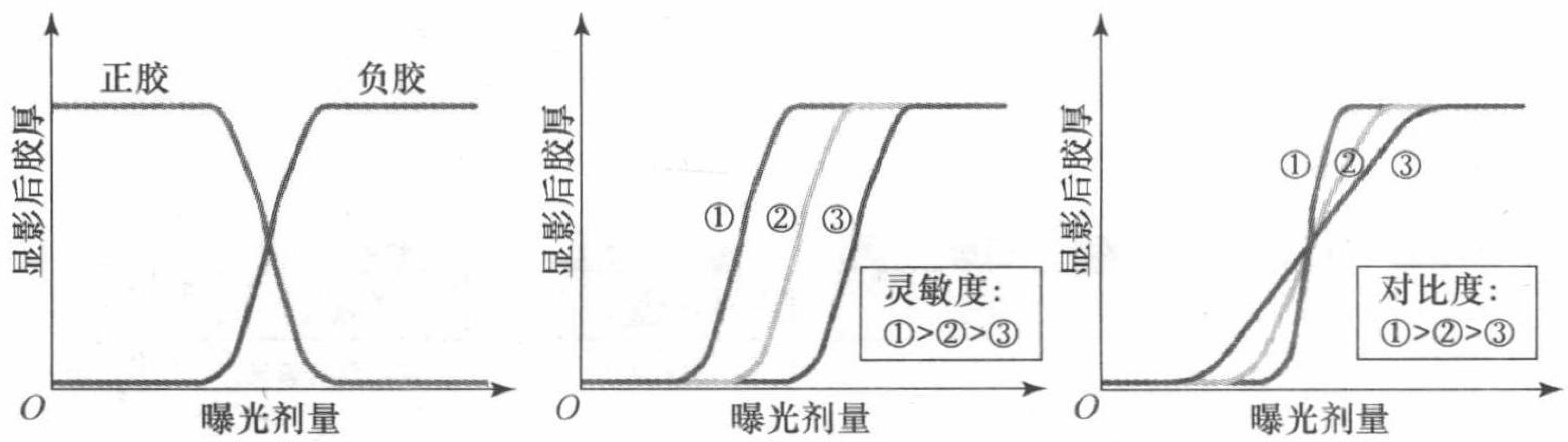


图 9.6 光刻胶的正负性、灵敏度和对比度

9.3.2 紫外光刻

典型的紫外光学曝光光刻技术包括接触式光刻、接近式光刻和光学投影光刻三种技术,如图 9.7 所示.

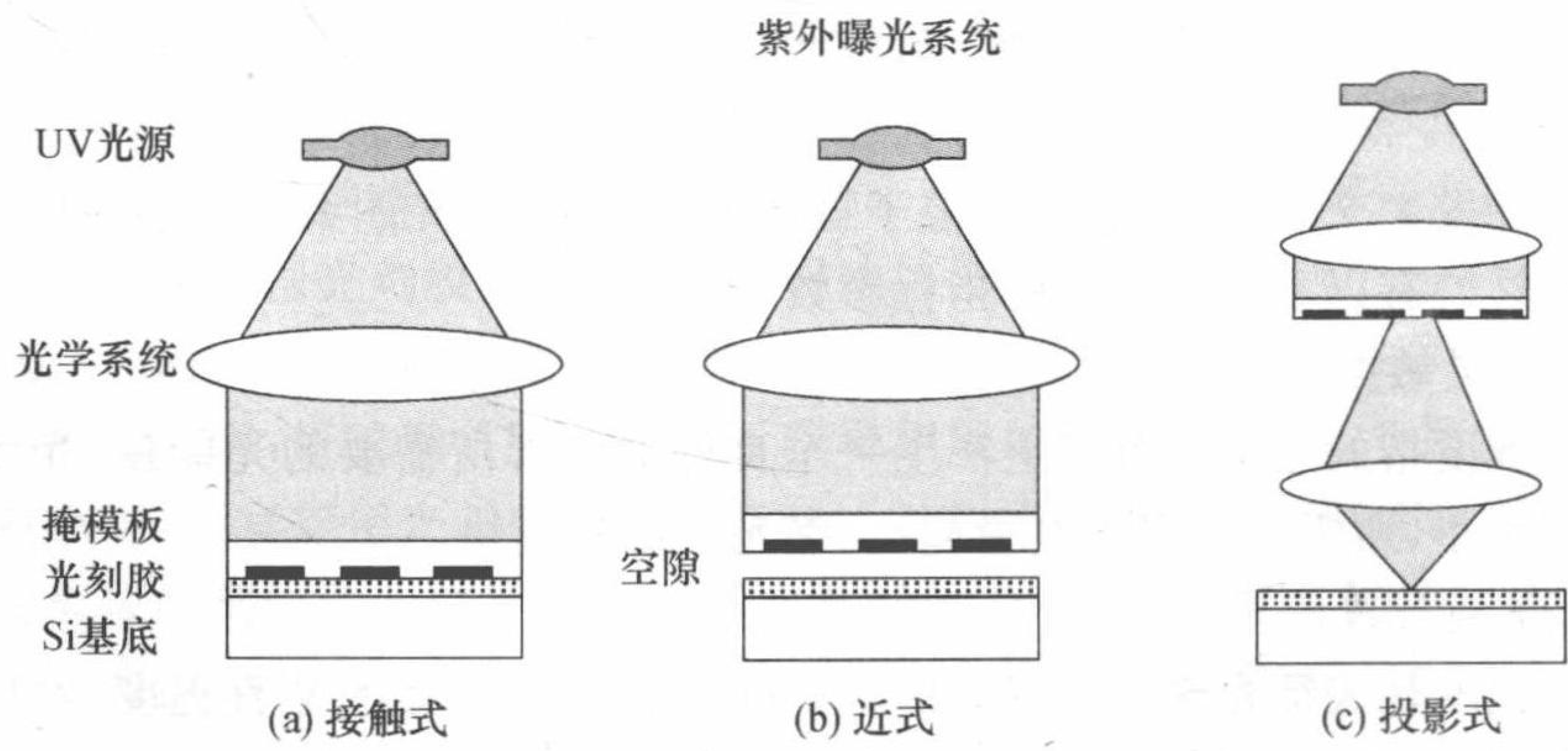


图 9.7 紫外光刻的三种曝光模式

接触式光刻技术是传统的光学光刻方式,是将预先制作好的掩模版直接和涂有光刻胶的硅片进行接触曝光.接触式光刻技术具有分辨力高、设备简单、操作方便和成本低等优点;但其致命的缺点是掩模和光刻胶直接接触,很容易损伤和玷污掩模版,因而掩模版的使用寿命短,多次使用将严重影响芯片的成品率.

接近式光刻技术是为了避免接触式光刻技术所产生的缺陷，在掩模和光刻胶之间有一定间隙的情况下进行曝光. 这种方式可以避免玷污和损伤掩模版，但衍射效应会使分辨力降低. 其分辨力经验公式为：$R=\sqrt{k\lambda h}$，式中 k 是表示光刻胶处理工艺的一个常数，λ 表示曝光波长，h 表示掩模与光刻胶之间的间隙.

光学投影光刻技术是利用光学投影成像的原理，通过投影物镜将掩模版上大规模集成电路图形成像在涂胶硅片上实现图形转移. 光学投影光刻技术能够得到接触式光刻技术的分辨力，而又避免接触式光刻技术容易损伤和玷污掩模版的弊端.

光学投影光刻技术是在接触式、接近式光刻技术基础上发展起来的. 采用投影光刻，可以延长掩模使用寿命，如采用缩小倍率投影物镜，还便于掩模制作. 光学投影光刻经历了分步重复光刻和步进扫描光刻的发展过程. 步进重复光刻是在光刻过程中使硅片台在 x，y 两个方向作步进运动. 步进扫描光刻是使硅片在一个方向作步进运动，另一个方向作连续扫描运动，在曝光过程中，需要掩模台与硅片台作与投影倍数相关的同步运动，即掩模台的运动速度是硅片台运动速度的投影倍数.

光学投影光刻是利用光学投影成像的原理，将掩模板上图形以分步重复（或步进扫描）曝光的方式将高分辨力图形转移到涂胶硅片上的光学曝光过程. 光学掩模版通常是镀有铬层的石英玻璃板，尺寸从 3～6 英寸不等，铬层根据电路图形被刻蚀成透光和不透光部分. 掩模板上的电路图形，经紫外光或深紫外光照明直接投影或缩小至图形的 1/4，1/5，1/10 后投影到硅片的抗蚀剂上. 经过显影步骤后，未被曝光部分的抗蚀剂被溶解（负性感光胶），或者被曝光部分的抗蚀剂被溶解（正性抗蚀剂）. 抗蚀剂被溶解部分的硅片裸露出来，然后进行下一步的刻蚀、扩散或金属沉积处理. 为了提高产量，硅片的直径越来越大，已经达到 12 英寸. 采用分步重复曝光方法，可将一组电路图形重复几十次转移到一片硅片上.

光学投影光刻是当今在超大规模集成电路制造生产线上应用最广、技术进步最快、生命力最强的光刻技术.

计算光刻分辨力的瑞利公式为

$$R = k_1\lambda/\mathrm{NA}, \tag{9.3.1}$$

式中 k_1 为工艺系数因子，λ 为曝光波长，NA 为投影光刻物镜数值孔径. 从上式可以看出，提高光刻分辨力可以通过缩短波长、降低工艺系数因子和提高投影光刻物镜数值孔径来实现. 缩短波长和提高投影光刻物镜数值孔径将对光学系统的设计、光学材料、光学镀膜、激光器光源、光刻工艺和环境条件等提出更高的要求. 对于大规模生产环境，工艺系数因子值通常为 0.7 左右. 对于实验室环境，则 k_1 值可降低到 0.5 左右.

根据瑞利公式，光刻分辨力可以通过缩短波长、降低工艺常数和提高投影光刻物镜的数值孔径来提高. 实践证明，缩短曝光波长是最为有效的途径. 光

学投影光刻技术自1978年诞生以来，先后经历了436 nm(g线)，365 nm(i线)，248 nm(KrF准分子激光)，193 nm(ArF准分子激光)等几个技术阶段.

除了缩短曝光波长外，不断减小工艺系数k_1也是进一步提高分辨力的十分重要的因素. 降低k_1值的途径包括改善照明条件、提高抗蚀剂性能、采用光学邻近效应校正和相移掩模等几个方面. 经过近10年的努力，已经使工艺系数因子k_1值在大生产环境中从0.7减小到了0.4. 上述因素更佳的组合将使k_1值减小到0.3乃至更小，这将成为今后一个时期光学光刻发展的一项战略措施. 当$k_1=0.25$时，便接近光学光刻的物理极限.

增大数值孔径也是提高光刻分辨力的重要途径、镜头的数值孔径已由最初的0.28，0.4，0.6逐渐增大到0.82. 最新推出的193 nm ArF光刻机的数值孔径最高已达0.85，分辨力提高到90 m，几乎到了极限. 由于传统的光学光刻机的工作环境是空气，其折射率$n=1$，因此光学镜头不可能实现NA>1的设计，使通过增大数值孔径来提高光刻分辨力的方法陷入困境. 然而，浸没式概念的引入使该方法焕发了蓬勃生机. 由于液体的折射率可以大于1，使得更大的NA可以实现.

表9.2列出了分辨率和数值孔径、曝光波长以及工艺系数因子之间的对应关系.

表9.2 分辨率和数值孔径、曝光波长以及工艺系数因子对应关系表

$R/\mu m$	NA	λ/nm	k_1
0.4	0.6	365	0.7
0.18	0.7	248	0.5
0.08	0.8	193	0.35
0.05	0.9	157	0.30
0.03	1.0	126	0.25

9.3.3 电子束光刻

1. 电子束曝光技术简介

电子束光刻(electron beam lithography)是利用电子束在涂有电子抗蚀剂(电子束光刻胶)的晶片上直接描画或投影复印图形的技术. 电子抗蚀剂是一种对电子敏感的高分子聚合物. 经过电子束扫描过的电子抗蚀剂发生分子链重组，使曝光图形部分的抗蚀剂发生化学性质改变. 然后经过显影和定影，获得高分辨率的抗蚀剂曝光图形. 电子束光刻技术的主要工艺过程为涂胶、前烘、电子束曝光、显影和坚膜. 现代的电子束光刻设备已经能够制作小于10nm的精细线条结构. 同时电子束光刻设备是制作光学掩膜版的重要工具.

电子束光刻与普通光刻技术最大的不同点是它利用电子束对电子抗蚀剂进行曝光，它不受光学衍射极限的制约，能够实现更高精度的曝光，通常用于曝光一些普通光刻无法完成的纳米结构.另外，电子束光刻系统是一种无掩膜直写系统，无需制作光刻板，版图通过计算机辅助软件完成后，直接可以曝光出相应的图形.这样的优点是具有很短的制造周期，省去了制造光刻板的时间，一个结构从设计出到最终的完成通常只需要几天的时间.而且，由于版图都是通过计算机辅助软件即可完成，对器件结构进行修改非常简单和方便，克服了普通光刻板无法进行修改的缺点，因此，它非常适合于科研.它的缺点是曝光速度很慢，它是一个逐点扫描的曝光系统，对于较大的结构需要很长的时间.虽然通过适当的增大电流，改善胶的特性等方法能相对的加快电子束的曝光速度，但是它的本质特性决定了电子束光刻不可能用于大规模的生产，一般只能用来制备一些纳米结构.

2. 电子束曝光系统

在这里我们对电子束光刻工作原理进行简单的介绍，参见图 9.8[2,3].电子束光刻设备产生高能量的电子束，通过电光系统(electrical-optical system, EOS)后聚焦在样片的表面，电光系统控制电子束光斑的聚焦，偏转，开关，从而实现电子束在样片表面进行扫描，曝光出需要的图形.EOS 是电子束光刻设备中最复杂和昂贵的部分，它控制着电子束的产生，聚焦，扫描等一系列活动.典型的电光系统包含以下部分：

(1) 电子枪.

电子枪是能产生稳定电子束的部件，通常由发射电子的阴极(cathode)、对电子束进行聚焦的栅极(suppressor)，以及提取电子束的阳极(anode)组成.

(2) 对准线圈.

我们可以发现，系统中有三个对准线圈.它们的作用是使电子严格沿着电光镜筒的中心向下传播.

(3) 开关光阑.

开关光阑是控制电子束通断的装置.通过控制电子束的通断从而实现逐点扫描，因此，该开关光阑的频率决定了电子束的扫描频率，是影响曝光速度的一个重要因素.它是利用静电场的工作原理，通过外加电场控制电子束的偏转.

(4) 可调节孔径光阑.

可调节孔径光阑是控制电子束通过多少的部件，它由多个孔径大小不同的光阑组成.它的作用是可间接的调节聚焦光斑的大小和电子束的电流，不同的应用可能选用不同的光阑孔径.一般来说，大的光阑具有大的光斑大小和大的电流，因此具有较快的曝光速度，写出来的线条也越平滑；小的光阑具有小的光斑大小和小的电流，可用来写更小的结构，但是曝光速度慢，信噪比较差，线条不如大光阑平滑.

(5) 电磁透镜.

电子在电磁场中受到力的作用而导致其运动轨迹发生改变，系统有多个电磁场透镜，其作用和光学系统中的透镜组一样，是将电子束聚焦在样品的表面. 如系统中电磁透镜二的作用是放大/缩小，物镜的作用是聚焦和进一步缩小光斑等.

(6) 电磁偏转器.

电磁偏转器(deflector)是用电磁场来控制电子束在样品表面进行扫描的装置，一般来说，这种偏转都会引入附加的像差，从而导致离中心远处的分辨率下降，实际上限制了电子束扫描的范围. 因此，一般采取小的写场，分步曝光的方法来获得整个样片比较好的曝光质量.

(7) 消像散单元.

消像散的作用主要是调节光斑的形状，使其接近理想的圆形. 为了获得高质量的曝光图形，这一点也是非常必要的.

此外，EOS 中经常还装有其他的一些装置，如二次电子探测器，它们用来收集从样片表面散射出来的电子，实现扫描电子显微镜(SEM)的功能.

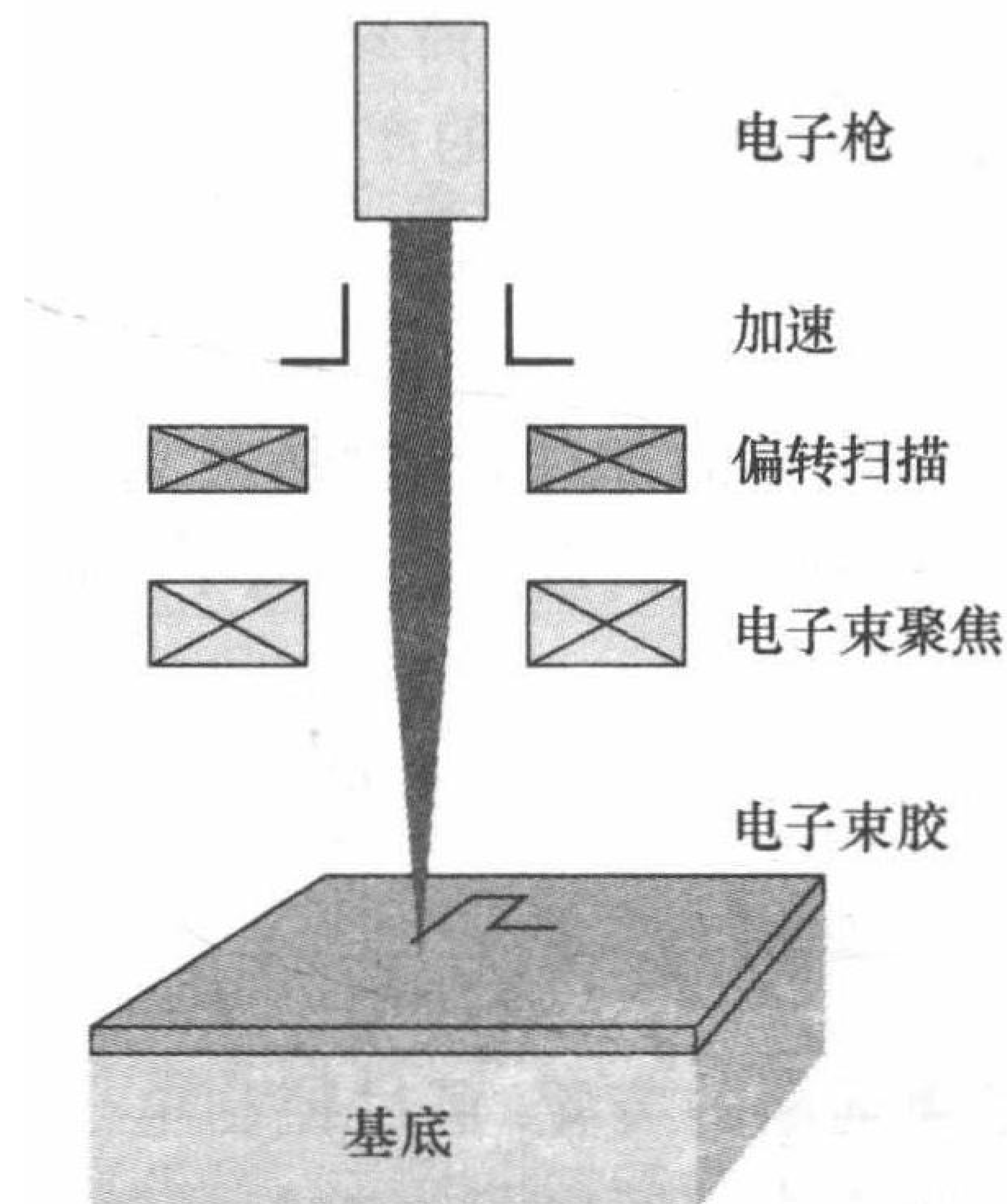


图 9.8 电子束曝光示意图

3. 电子束曝光过程中的邻近效应

影响电子束曝光精度的一个很重要的因素是邻近效应. 电子束打在涂有光刻胶的样片表面后，电子在光刻胶中会发生散射(前向散射)，同时，部分射入基底的电子会反射回来(背向散射)重新进入光刻胶中(参见图 9.9)，这些因素使得我们希望被曝光区域附近的光刻胶也会发生不同程度的曝光，从而成为限制电子束曝光分辨率的主要因素之一，我们称之为邻近效应(proximity

effect)[4,5].

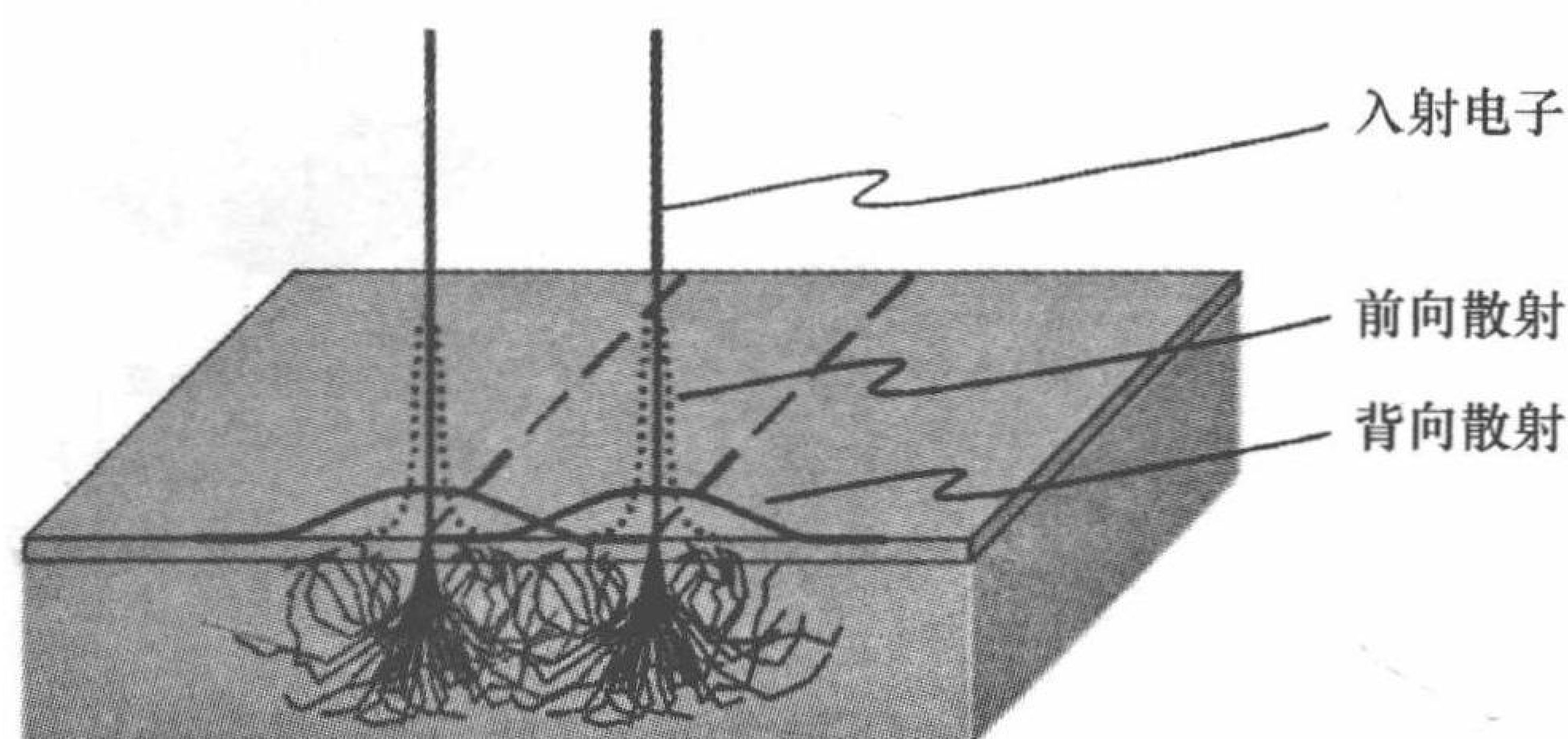


图 9.9　电子束在电子束光刻胶中的前向散射和背向散射

人们一般引入双高斯函数来表示电子束单点曝光时在光刻胶中的能量分布：

$$f(r) = \frac{1}{\pi(1+\beta)}\left\{\frac{1}{\alpha^2}\exp\left[-\left(\frac{r}{\alpha}\right)^2\right]+\frac{\eta}{\beta^2}\exp\left[-\left(\frac{r}{\beta}\right)^2\right]\right\}, \quad (9.3.2)$$

式中 α,β,η 为曲线拟合函数. 其物理意义是：α 和 β 分别表征前向散射和背向散射的范围，实际上就是图 9.9 所示高斯函数的半宽度，单位为 μm. η 表示背向散射和前向散射的能量之比. 表 9.3 给出了硅基底上 0.5 μm 厚电子束胶中这些参数的典型值. α 是理论计算得到的，β 和 η 是实验得到的，方括号中数据是外推的结果. 为了更好地反映电子束在电子束胶中的能量分布，有人在双高斯函数的基础上添加一个指数项更好地拟合实验结果，修正后的拟合公式如下[6]：

$$f(r)=\frac{1}{\pi(1+\beta+\nu)}\left\{\frac{1}{\alpha^2}\exp\left[-\left(\frac{r}{\alpha}\right)^2\right]+\frac{\eta}{\beta^2}\exp\left[-\left(\frac{r}{\beta}\right)^2\right]+\frac{\nu}{2\gamma^2}\exp\left[-\left(\frac{r}{\gamma}\right)^2\right]\right\}, \quad (9.3.3)$$

其中 γ 是指数函数的衰减函数，ν 是指数项强度于前向散射强度之比.

表 9.3　不同加速电压下，α,β,η 的典型值[7]

加速电压/kV	α/μm	β/μm	η/μm
5	1.33	[0.18]	[0.74]
10	0.39	[0.60]	[0.74]
20	0.12	2.0	0.74
50	0.024	9.5	0.74
100	0.007	31.2	0.74

注：方括号表示数据是外推法得出来的，而不是实验测得的.

根据上表中的参数以及双高斯近似，我们得出不同加速电压下电子束在胶中的能量分布如图 9.10 所示.

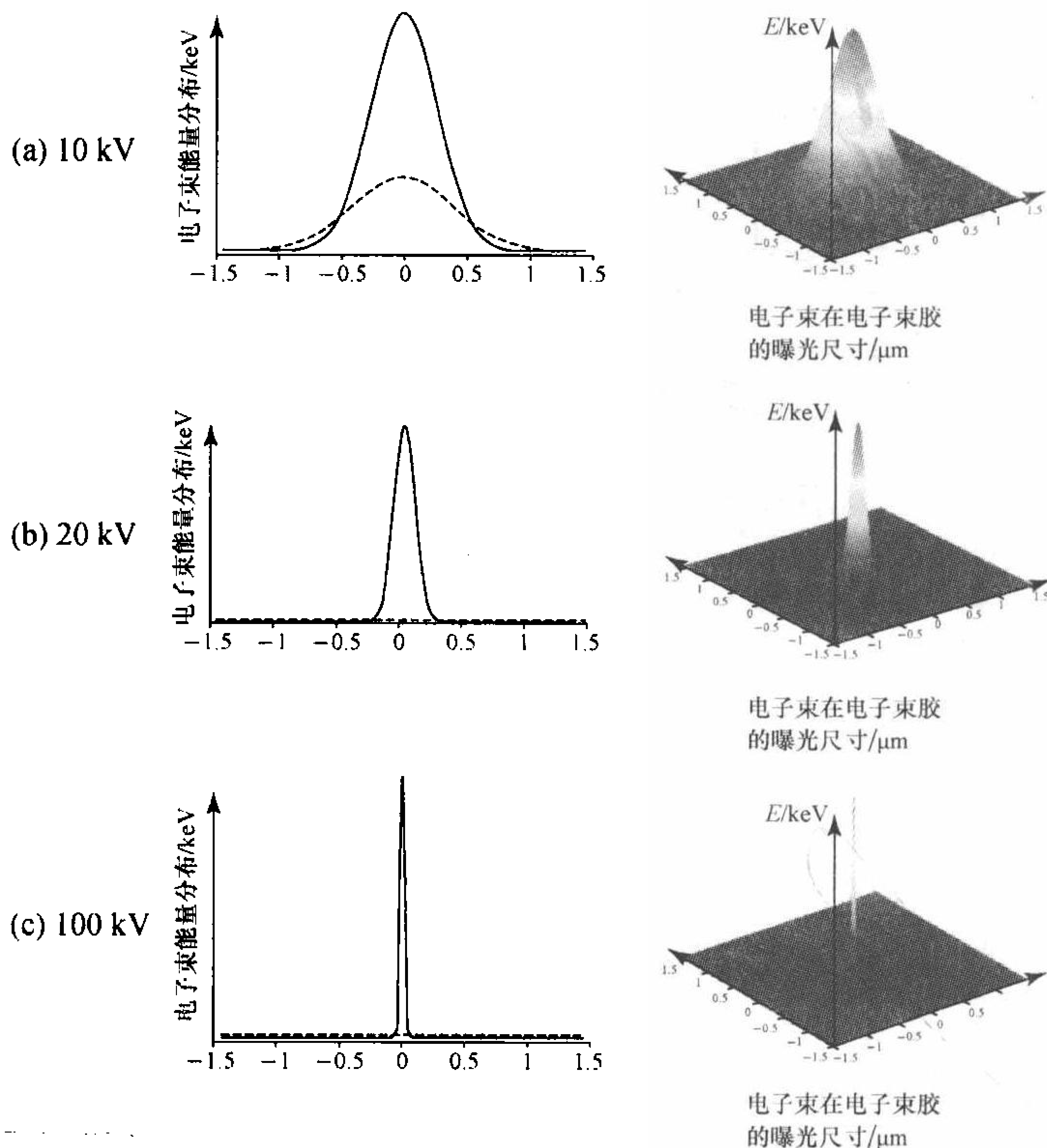


图 9.10 不同电压下的邻近效应双高斯分布.实线表示前向散射能量分布,虚线表示背向散射能量分布

更方便描述邻近效应对曝光精度影响的公式是调制传递函数(modulation transfer function,MTF),将式(9.3.2)作傅里叶变换和归一化处理后,我们可以得到传递函数的表达式

$$M=\frac{1}{1+\eta}\left[\exp\left(-\frac{\pi^2\alpha^2}{p^2}\right)+\eta\exp\left(-\frac{\pi^2\beta^2}{p^2}\right)\right], \tag{9.3.4}$$

其中 p 为图形的空间频率(对光栅来说 $p=1/T$),实际表示了图形的大小和密集程度,因此式(9.3.4)可以很方便地表示邻近效应对不同大小曝光图形的影响,从而进行修正.理想情况下,对于所有 p,M 都应等 1,表示没用任何图形信息丢失.当存在电子散射时,$M<1$ 而且依赖于 p.M 越小表示邻近效应越明显,对曝光结果影响越大.

从图 9.11 可以看出,高的加速电压下(100 kV)的曲线有较长的平坦部分,说明采用高能量的电子束有更大的工艺窗口,可以同时曝光高密度和低密度的图形,尤其在高密度图形的曝光方面比低电压有优势.但高能量电子束会导致电子束胶的灵敏度降低(需要相对较长的曝光时间),而且会导致基底加热,对某些基底还有损伤,实际曝光中需要综合考虑.

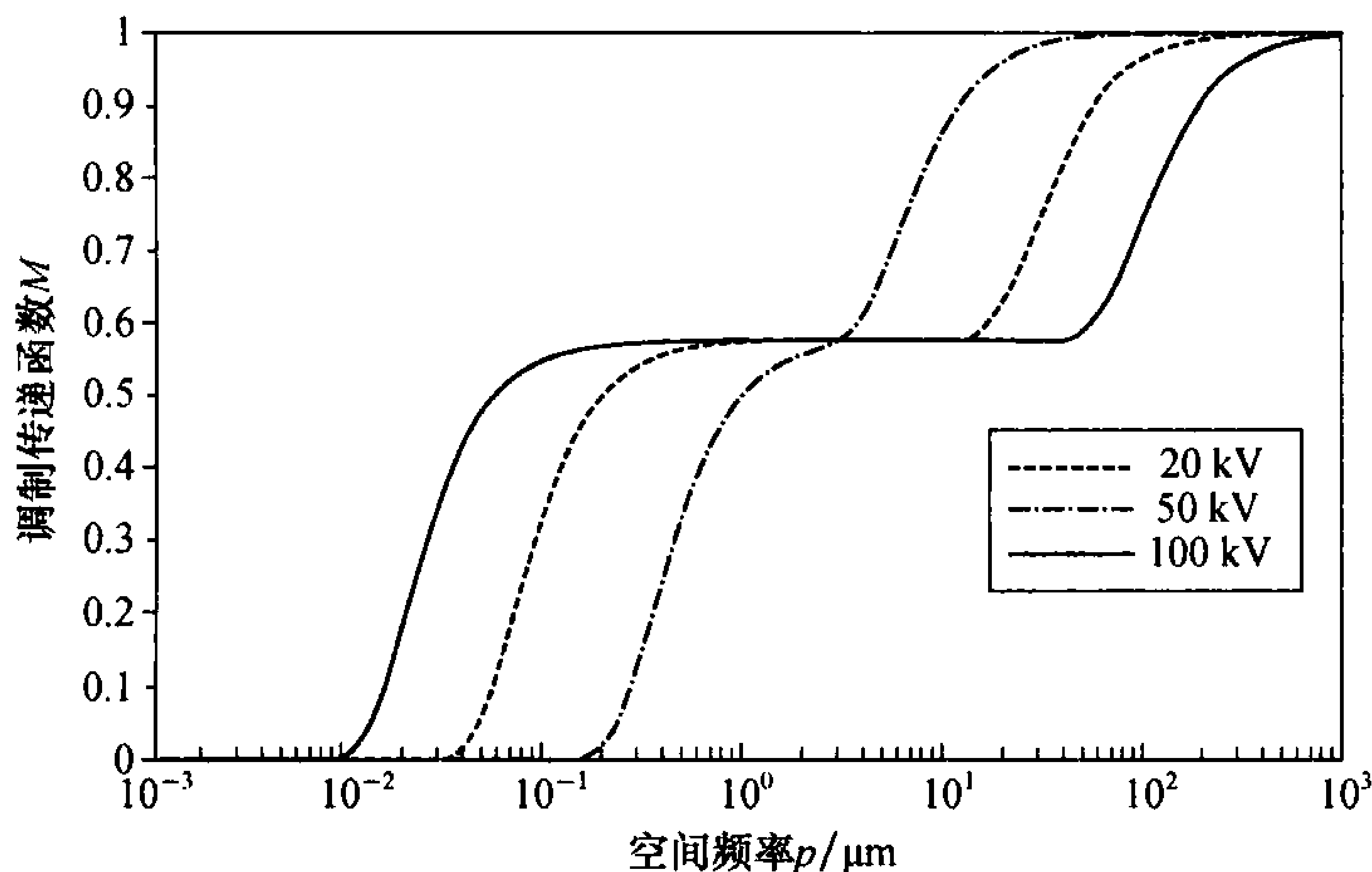


图 9.11　不同加速电压下的调制传递函数曲线图

从上面分析可以看出，前向散射的结果主要是导致线条的展宽，背向散射会导致图形对比度变差，区域曝光的中心部分比边缘曝光充分. 通常我们采用以下一些方法来减小邻近效应的影响：

① 采用高的加速电压(减小前向散射)或低的加速电压(减小背向散射).

② 采用多层胶工艺.

③ 选用高对比度的电子束光刻胶.

④ 选用低原子量的材料作为基底，减小背向散射.

⑤ 使用薄膜基底，让电子束透射出薄膜，消除背向散射.

上面的一些方法虽然可以一定程度上减小邻近效应的影响，但实际工艺中很多条件无法满足，如：基底材料无法改变、可利用的电子束胶种类受限等. 人们希望有一些不依赖于设备和工艺条件的减小邻近效应的方法[8]. 一般有以下几种方法可以对邻近效应进行修正：

⑥ 剂量校正. 剂量校正方法就是考虑图形和图形间的相互作用，对图形的总曝光剂量进行计算，然后为图形中每个像素点或最小单元精确地指定所需剂量，使得图形中各个地方的总剂量一致的方法. 如果电子束的能量分布模型正确，该种方法进行剂量校正可以得到比较精确的结果，但是对于复杂的图形计算却需要大量的时间.

⑦ 图形补偿校正. 图形补偿校正是通过改变图形尺寸来补偿因电子散射而造成的图形畸变的校正方法. 在这种方法中，拐角部分通过补偿图形可获得较好的直角，密集图形通过减小尺寸来获得剂量修正.

⑧ 背景曝光修正[9]. 该方法又叫 GHOST 方法，是通过散焦的电子束对曝光图形的反图进行曝光，从而拉平背景剂量实现邻近效应的修正. 这种方法的优点是无需大量校正计算，缺点是总曝光数据量增大和时间增长，并且降低了曝光

图形的对比度,故一般很少采用.

4. 电子束曝光剂量的选择

最佳的曝光剂量跟很多因素有关,如图形的大小、密度、电子束胶的厚度、加速电压、基底材料等.由于每一项参数的可变性都很大,我们很难建立一套精确的模型来表征所有的参量,只能根据一些基本原则来选择曝光剂量.

一般来说,较大的图形尺寸需要较小的剂量,这是因为在大的曝光区域邻近效应作用的范围更大,每点实际获得的曝光剂量也较大,所以一般用较小的剂量.同理,当图形的密度增大时,也要相应减小曝光剂量.对于较厚的胶,显然需要比较大的剂量来获得足够的曝光.另外一个重要因素是加速电压,大的加速电压意味着小的电子束斑,同时意味着散射减少,因此需要增大曝光剂量.曝光剂量的选择还跟基底材料有关,基底材料的原子量较大时,电子的反射也越厉害,因此曝光剂量需要相应减小.

表 9.4 表示了最佳曝光剂量随各参数变化而变化的关系,在实验中,往往先要用一系列不同的剂量来曝光需要的图形,然后从中选取最合适的曝光剂量.如图 9.12 所示,在同一样片上利用不同曝光剂量曝光 ZEP 电子束胶[10]上相同图形,测得图形尺寸的变化.

表 9.4 曝光剂量与各参数间的关系

参量	变化	最佳剂量
图形尺寸	↑	↓
图形密度	↑	↓
电子束胶厚度	↑	↑
加速电压	↑	↑
基底材料原子质量(atmic mass unit, AMU)	↑	↓

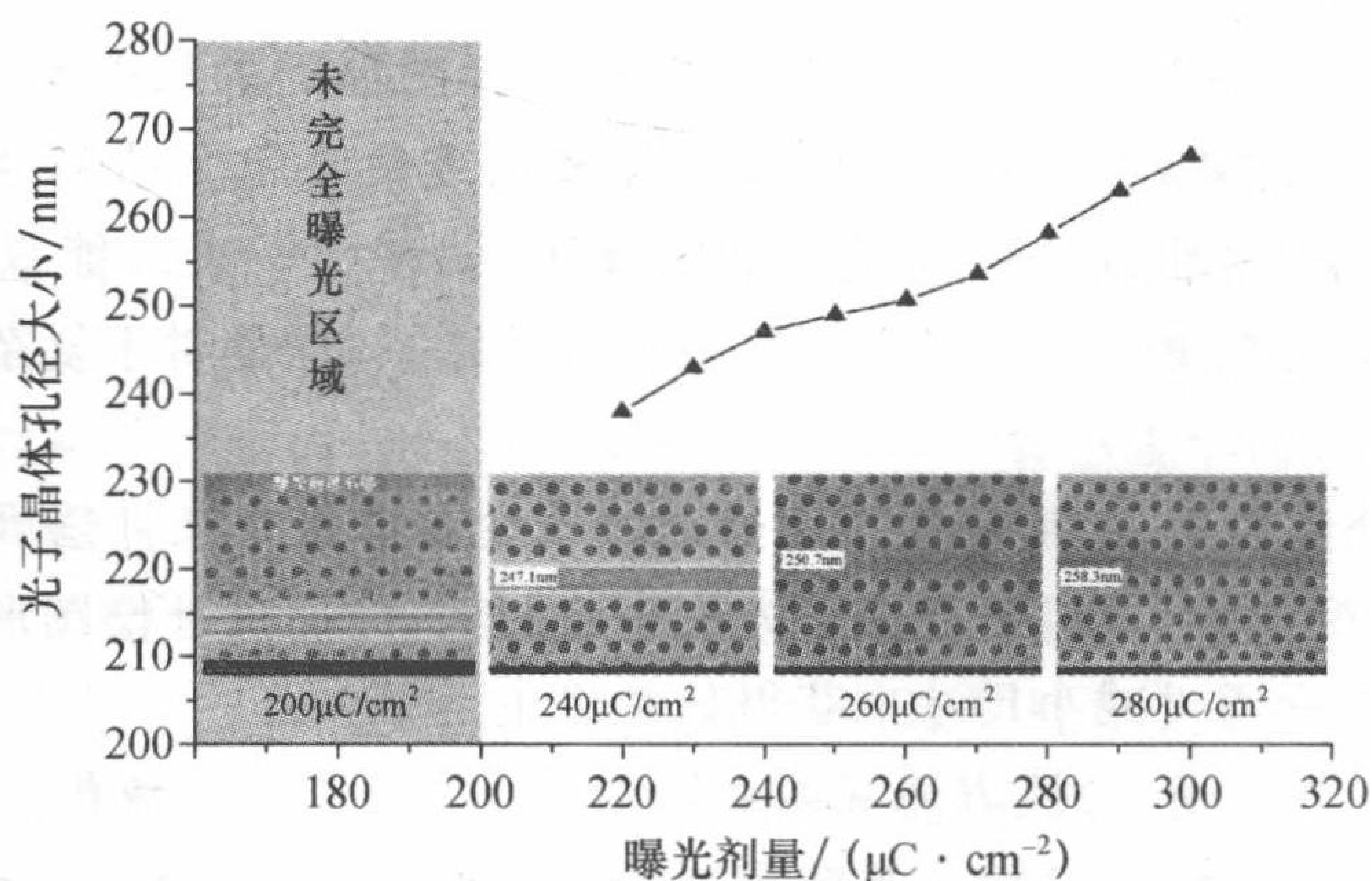


图 9.12 同一样片上不同曝光剂量造成的光子晶体孔径大小变化

9.3.4　混合光刻

紫外光刻曝光速度快但精度低，电子束曝光精度高但速度慢. 实际的器件一般同时包含较大的结构和很小的结构. 因此，在实际应用中，人们往往结合两种曝光方法来实现光刻，我们称之为混合光刻技术.

混合光刻流程一般如下：① 紫外曝光/电子束曝光制备对准标记，当用紫外曝光制备对准标记时，有时可以第二步同时进行；② 紫外曝光较大的结构，并刻蚀；③ 电子束对准曝光小的结构，刻蚀. 有时也可以先用电子束曝光获得小的结构和对准标记，然后利用紫外光刻对准曝光大的结构. 但是，紫外光刻对准精度不高，通常在几微米，而电子束曝光的对准误差仅为几十纳米. 因此，一般情况下都是采用先紫外光刻，后电子束套刻的方法.

混合光刻最关键的在于前后两次光刻的套刻. 同紫外光刻用光学显微镜对准不同，电子束光刻只能用电子在对准标记上的扫描信号来判断标记的位置. 通过在标记上扫描反射信号的变化来确定标记的位置，因此电子束曝光的对准精度与标记本身的好坏有很大关系. 好的标记需要边缘光滑，而且在电子束扫描下有足够的信噪比. 图 9.13 为电子束扫面对准标记获得坐标的工作示意图.

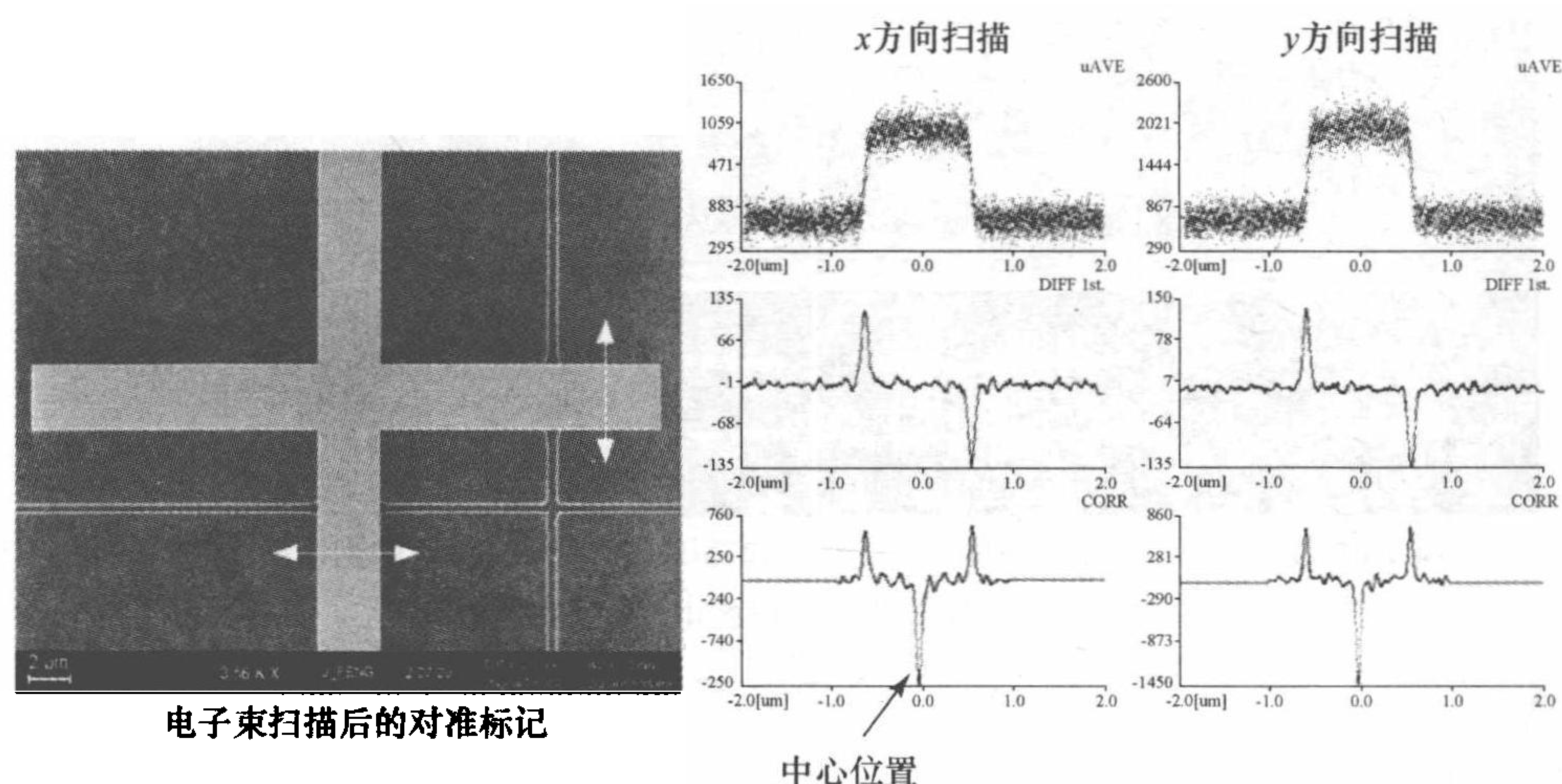


图 9.13　电子束对准标记的扫描和位置的判定

一般来说，标记的宽度约为 3 μm，可以由紫外曝光或电子束曝光来获得，但实验表明，同一电子束曝光系统制作的对准标记是最好的. 按照材料来划分，标记有两种：

（1）刻蚀标记.

为了获得足够高的信噪比，对于硅来说，刻蚀的标记一般至少需要 1 μm 以上，较理想的深度是 2 μm. 而 SOI 硅片要获得 2 μm 的刻蚀深度，需要三次刻蚀，

分别刻蚀顶层 Si，埋层 SiO_2 和 Si 基底，获得如此深的边缘光滑的刻蚀标记有一些困难. 图 9.14 为 SOI 上制备的刻蚀标记的侧面 SEM 照片，可以发现侧壁的光滑性不好，这会影响电子束扫描的信号.

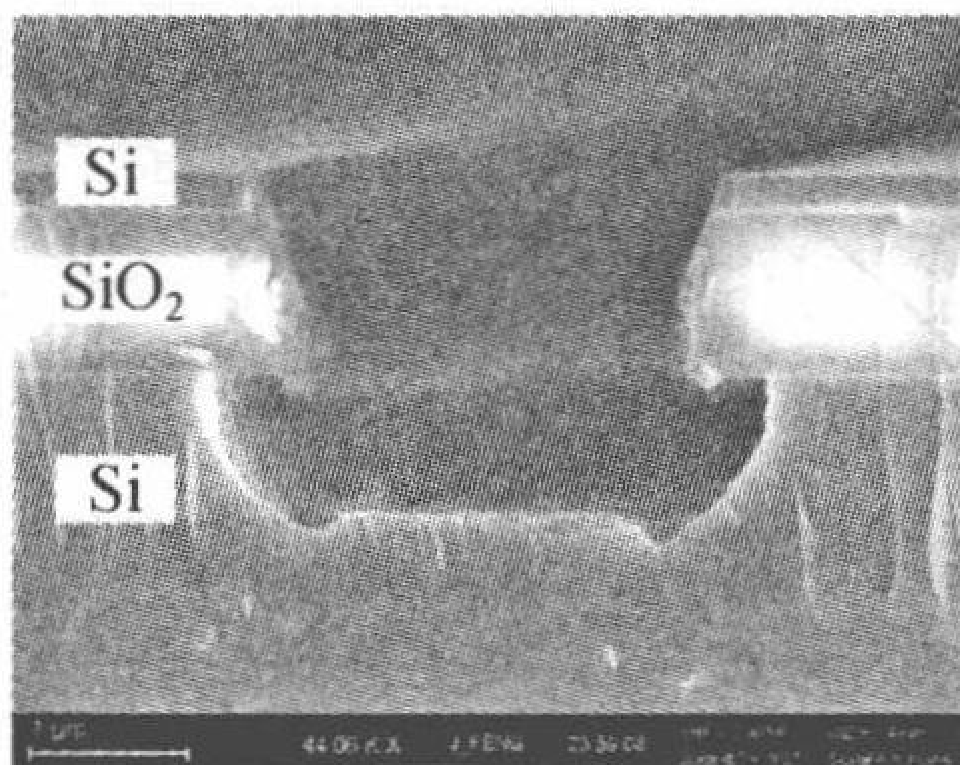


图 9.14 SOI 上刻蚀对准标记的侧面 SEM 照片

(2) 金属标记.

金属标记一般由剥离工艺获得，厚度一般需要 50 nm 以上，比较理想的厚度是 100 nm. 金属材料的选择也很重要，最好选用重金属(如 Au，Cr 等)，以获得好的信噪比. 我们一般用金作为对准标记的材料，但是由于金与硅的粘附性很差，一般先沉积 10 nm 的 Ti，然后沉积 90 nm 的 Au.

另外，有些后续工艺，如薄膜沉积、刻蚀等，可能会损坏对准标记，我们在进行这些工艺时需要提供一些保护措施. 我们对不同的对准标记进行了对比. 图 9.15 所示为不同对准标记的电子束扫描信号. 我们发现，Au 的信噪比最好，其次是深刻蚀的 Si. 轻金属 Al 作为对准标记，信噪比很差. 图 9.15(d)所示为未保护好的对准标记，在电子束扫描下基本无法分辨.

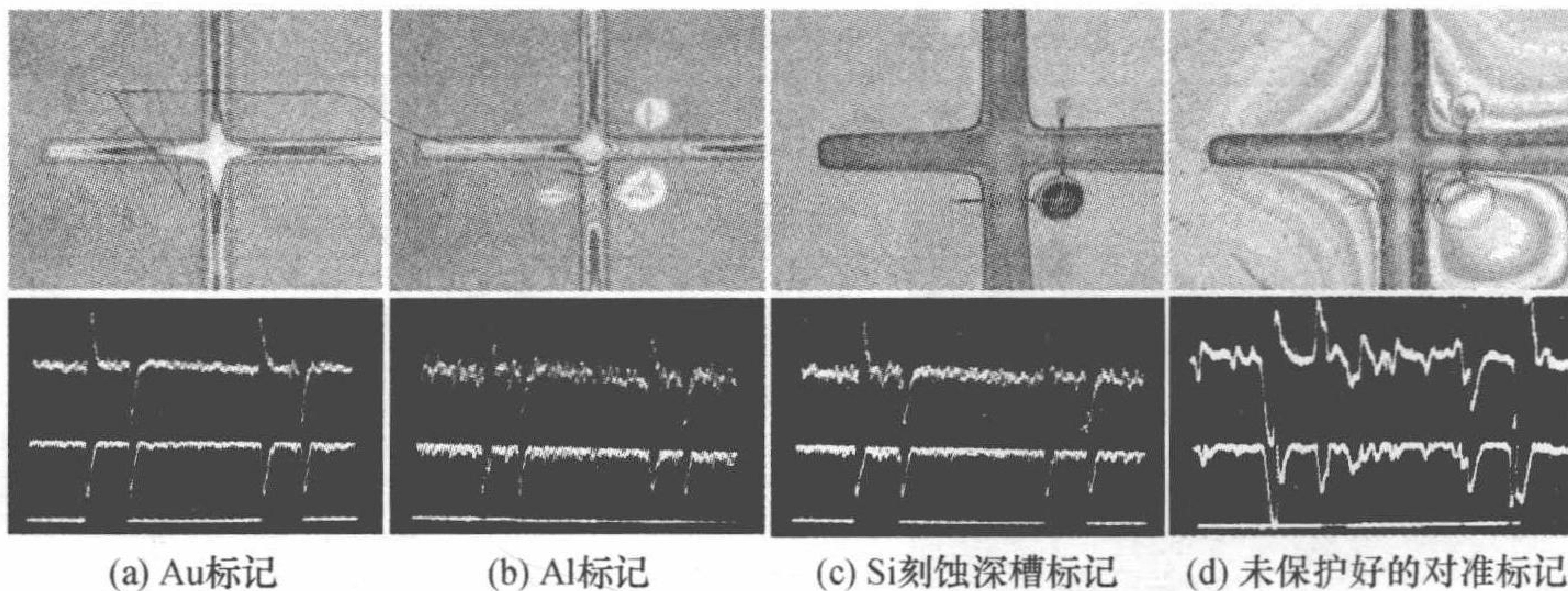


图 9.15 不同对准标记的电子束扫描信号

§9.4 刻蚀技术

9.4.1 刻蚀技术简介

刻蚀技术是图形转移的重要技术之一，所谓刻蚀就是利用化学或者物理方法将图形从掩膜转移到材料中的过程. 图 9.16 所示为常用刻蚀方法的分类，由于湿法刻蚀可控性差，小尺寸图形刻蚀困难，而且很难获得各向异性的刻蚀，我们主要研究干法刻蚀的方法.

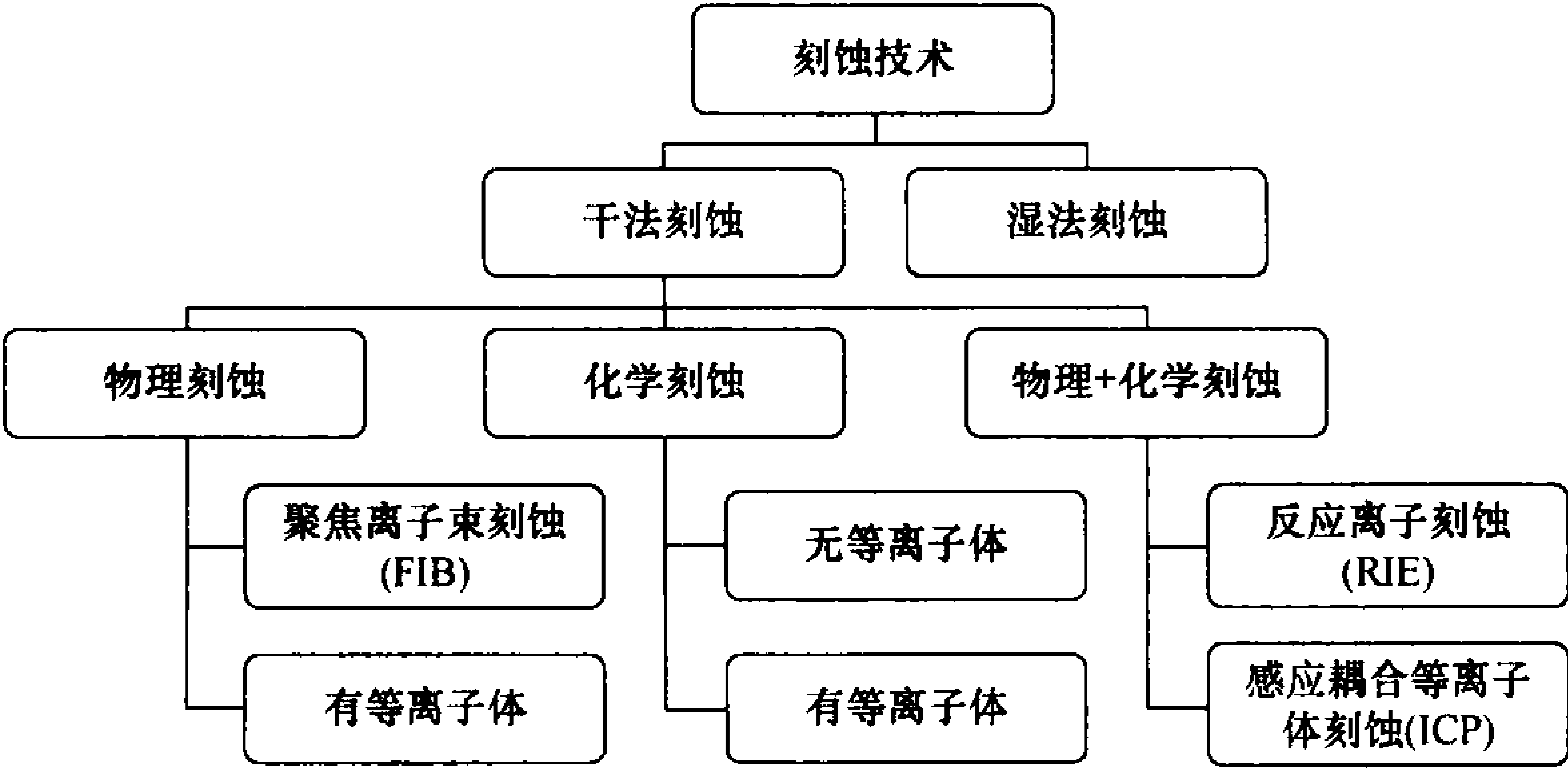


图 9.16　常用刻蚀技术分类

干法刻蚀又可以分为三类：物理刻蚀，如聚焦离子束（focused ion beam，FIB）；化学刻蚀，如 XeF_2 刻蚀硅；物理＋化学刻蚀，如反应离子刻蚀（reactive ion etching，RIE）和感应耦合等离子体（inductively coupled plasma，ICP）。物理刻蚀的原理是用离子轰击样片表面，打出材料分子，其优点是各向异性刻蚀，可以得到很垂直的侧壁形貌，缺点是选择比差，样片表面有损伤；化学刻蚀的原理是利用活性分子与材料发生化学反应，其优点是选择比强，缺点是很难获得垂直的侧壁形貌；物理＋化学刻蚀的方法是结合两种刻蚀的机理，不但可以获得好的刻蚀方向性和选择性，而且大大加快了刻蚀速度，成为目前主流的刻蚀方法。图 9.17 显示了三种刻蚀方法的机理和刻蚀速率对照关系。

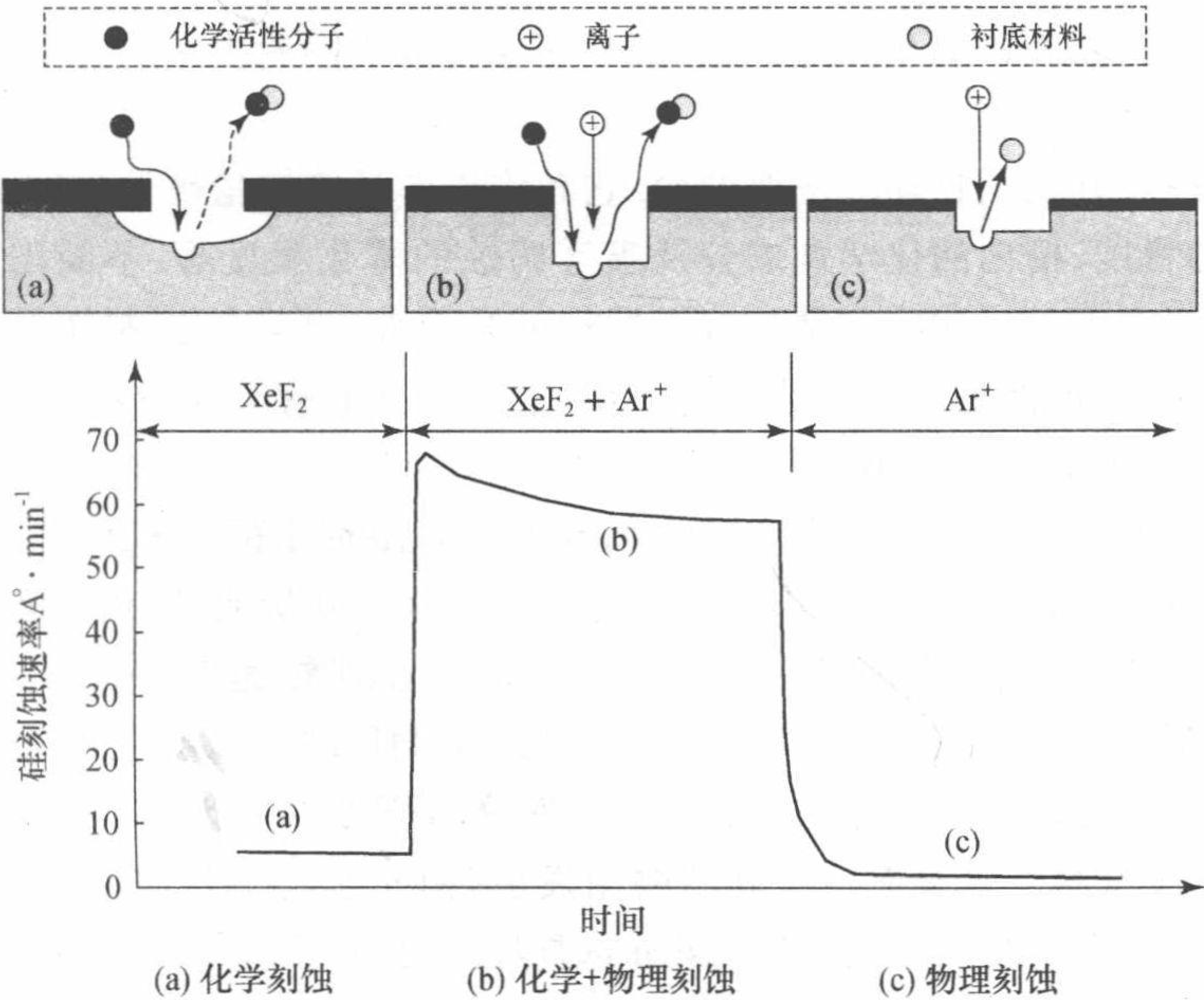


图 9.17　三种刻蚀方法的原理图以及刻蚀速率曲线

刻蚀过程中，我们主要关注的参数有以下一些：

① 刻蚀速率，是指刻蚀过程中除去被刻蚀层的速度.

② 刻蚀选择比，是指刻蚀过程中去除被刻蚀材料的速度和去除掩膜的速度之比.我们希望刻蚀选择比越高越好.

③ 刻蚀方向性，有时也叫刻蚀陡直度，它反应了刻蚀横截面侧壁的陡直性.有时我们希望陡直性越高越好(各向异性刻蚀)，有时会希望各向同性的刻蚀.

④ 刻蚀粗糙度，是指刻蚀造成的结构侧壁不光滑的程度.

⑤ 刻蚀损伤，是指刻蚀过程中由于离子的轰击对材料基底损伤的大小.

9.4.2 湿法刻蚀

湿法刻蚀是一个纯粹的化学反应过程，有着缺乏各向异性、差的工艺控制和过度的颗粒玷污等严重的缺点.然而湿法刻蚀有着高的选择比，通常不产生基底损伤.结果是，尽管这种工艺已经不如以往那么普遍使用，它还继续广泛地被应用于非关键尺寸的任务中.

由于反应产物通常出现在刻蚀溶剂中，故湿法刻蚀由三个步骤组成：① 刻蚀剂移动到基底表面；② 与暴露的膜发生化学反应生成可溶解的副产物；③ 从基底表面移去反应生成物.由于所有这三个步骤都必须发生，最慢的一个步骤称为速率限制，它决定着刻蚀速率.由于通常我们都想获得高的均匀的，受控良好的刻蚀速率，所以湿法刻蚀溶剂通常被以某种方式搅动，用以帮助刻蚀剂到达硅片表面，并且帮助去除刻蚀生成物.一些化学刻蚀工艺使用一种不间断的酸性喷雾以确保一直有新鲜的刻蚀剂供应，但是这样导致了随着大量的化学试剂浪费，生产成本上升.

现在已经开发了自动湿法腐蚀台.它允许操作者能够精确地控制腐蚀时间、槽温、搅动程度、槽内的化学剂成分及喷雾腐蚀的雾化程度等.不断增加使用甚至是在热状态下过滤这些非常具有侵蚀性的化合物，已经给控制相关的颗粒沉积提供了帮助.即便是随着这些改进，但对于大部分的特征尺寸小于 2 μm 的工艺中，湿法腐蚀还是不实用的.

最常见的湿法腐蚀工艺之一是在稀释的 HF 溶剂中的二氧化硅湿法腐蚀.常用腐蚀液配比是 6∶1,10∶1,20∶1.意味着 6 份、10 份或 20 份体积的水与一份 HF 混合.6∶1 的 HF 腐蚀液腐蚀热氧化 SiO_2 的腐蚀速率是 120 nm/min，而采用 PECVD 沉积的氧化层腐蚀速率则更快.在 HF 之中沉积膜的腐蚀速率与热氧化膜层的速率比通常被用来测量膜的密度.HP 腐蚀液对 SiO_2 与 Si 有着极高的选择性.这其中也会有一些硅的腐蚀发生，由于水将缓慢氧化硅表面，然后 HF 将该氧化层腐蚀.HF 对 SiO_2 与 Si 选择能力通常好于 100∶1.然而，氧化层在溶剂中的湿法腐蚀，完全是各向同性的.

9.4.3　干法刻蚀

上一小节中我们说到，物理＋化学刻蚀的方法是目前主要的刻蚀方法，其最典型的系统是反应离子刻蚀(RIE). 感应耦合等离子体(ICP)刻蚀看做是改进型的 RIE. 它们的工作原理大致相同，因此我们经常把它们拿来一起讨论.

1. RIE 和 ICP 系统工作原理

图 9.18 所示为典型的 RIE 和 ICP 系统，ICP 相比 RIE 最大的区别就是多了一个 RF 电源. 它们的工作原理都是利用 RF 电源在真空腔内产生等离子体，利用等离子体与被刻蚀样片发生化学和物理反应，气态副产物通过真空泵被抽出腔室.

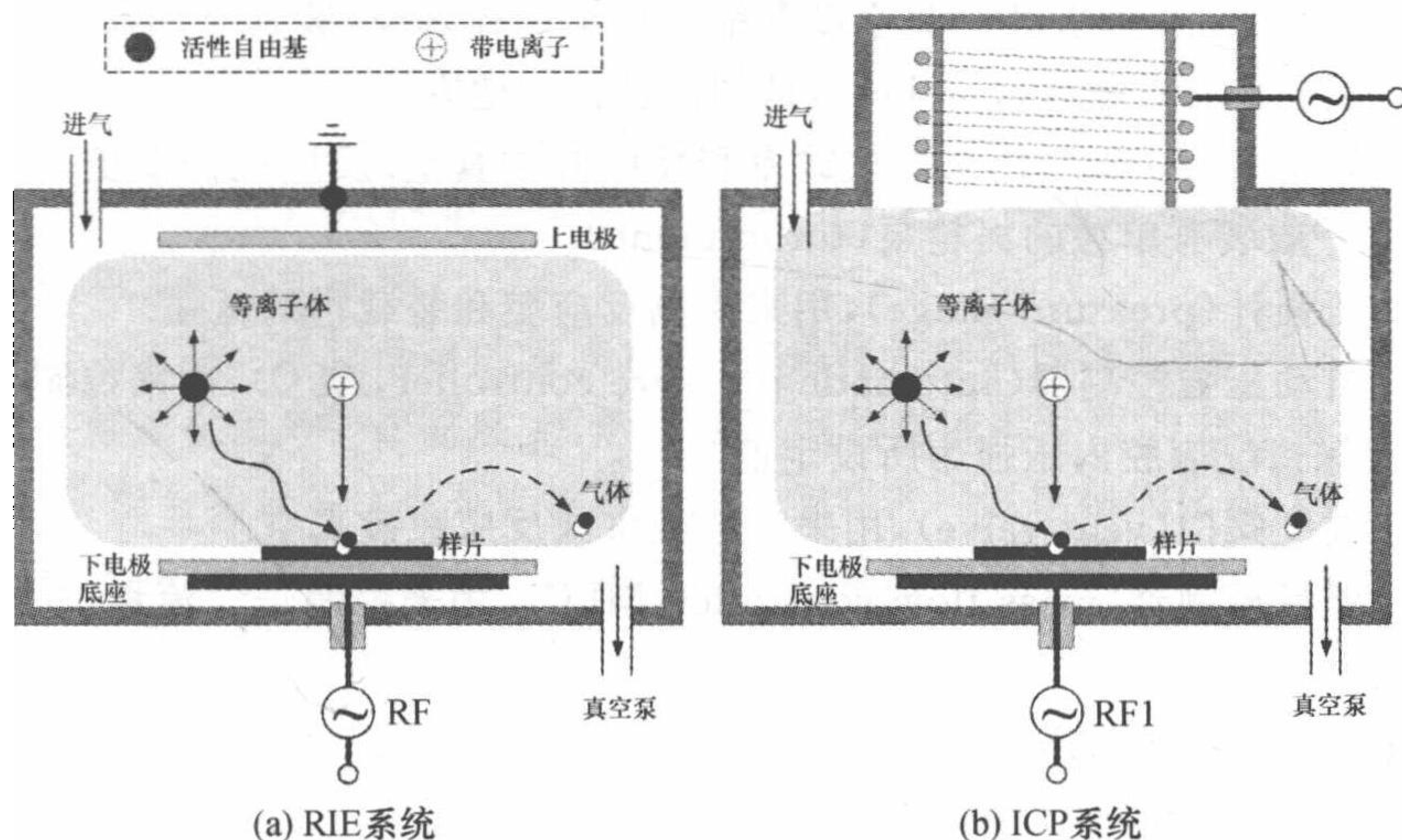


图 9.18　RIE 系统和 ICP 系统基本结构示意图

等离子体是一系列带电原子或原子团，电中性分子或自由基团，自由电子等组成的集合，宏观上呈电中性. 等离子体一般在真空条件下由气体放电产生，实际参与刻蚀的只是其中的“活性自由基”和“带电离子”，如图 9.18 所示. 它们在等离子体中占据的比例很小，一般活性自由基约为 2%，离子和电子总共约为 0.001%.

选用 RF 电源而不采用直流(direct cwrrent, DC)电源，是因为多数待刻蚀材料为绝缘体或半导体，直流偏压导致的电荷积累所产生的内部电场会很快中和外加的直流电压，使等离子体不能持续. 因此，通常只有在材料为导电金属时才用直流电压来产生等离子体(如 DC sputter). RF 频率选择大于等离子体的放电频率，从而维持系统等离子体的连续，一般采用 13.56 MHz.

我们注意到在 RIE 系统中，上电极经常与腔体相连并且接地. 这样做的目

的是增大上电极的有效面积，获得对下电极的正向偏压，从而带正电的离子在该偏压下轰击样片表面. 而在沉积系统(如 PECVD)中，为了获得在样片表面的沉积，通常采用相反的构造，增大下电极面积.

由于反应的副产物必须是气态，否则会附着在样片表面阻止刻蚀的继续进行. 而有些产物的气化温度较高，此时往往通过底座给样片加温，使刻蚀顺利进行. 另外，有些刻蚀希望降低样片的表面温度来获得好的刻蚀形貌，此时就需要通过底座给样片降温(如氦气冷却或水冷等).

2. ICP 系统

一台 ICP 设备一般由以下部分组成：

① 装载室(loadlock)，用于传递样片至真空腔或从真空腔中取出样片. 它的作用是更好保持刻蚀腔体的真空度和缩短进出抽样真空(破真空)的时间.

② 反应腔(process chamber)，是刻蚀进行的地方.

③ 真空泵(pump)，用来对装载室和反应腔抽真空. 反应腔需要更高的真空度，因此一般装有单独的涡轮泵(turbo pump).

④ 压强计(pressure gauge)，用来探测反应腔和装载室的压强.

⑤ 自动压强控制器(automatic pressure controller, APC)，用来控制抽气口的大小，保持反应腔内压强为所设定值.

⑥ 放气阀门(vent valve)，用来对装载室或反应腔破真空.

⑦ 流量控制器(mass flow controller, MFC)，用来控制进气流量. 一般装于进气线路中.

⑧ 射频发生器(RF generator)，用来产生射频信号.

3. ICP 与 RIE 相比的优点

从上面的讨论我们知道，ICP 和 RIE 的主要区别就是多了一个射频源. 该射频源的作用是在腔室顶部产生一个磁场，带电粒子在磁场内运动轨迹为曲线，有更多的机会发生碰撞而产生等离子体，因此它是一种高密度等离子体(high density plasma, HDP)的方法. 该射频源的引入，大大加强了刻蚀的可控性和刻蚀性能，总结起来主要有以下几点：

① 更高的离子密度(RIE 为 $10^9 \sim 10^{10}/cm^3$, ICP $> 10^{11}/cm^3$)和自由基密度，意味着更高的刻蚀速率；

② 工作在更低的压强(一般 RIE 在 100 mTorr 左右，ICP 在 10 mTorr 左右)，意味着更好的刻蚀形貌；

③ 更低的偏压 DC 偏压(RIE 一般大于 100 V, ICP 可控制在很小)，意味着可以提高刻蚀选择比和减小基底损伤；

④ 分离了等离子体产生电压和刻蚀偏压，增强刻蚀灵活性，且可工作为 RIE 模式.

§9.5 系统集成

9.5.1 系统集成方法

作为光电子器件的材料，硅未必会是最好的和唯一的选择，正如我们前面章节中所讨论的，在某些方面硅存在本质上的缺陷. 但它的一个最大优点是成熟的硅基微纳加工技术和可以实现大规模的集成. 如同电子器件中 CMOS 晶体管一样，众所周知 CMOS 晶体管的性能不是最好的，但这种晶体管的可大规模单片集成特性使得它成为人们普遍的选择. 人们现在已经能将许多光学器件集成在一个基底上，实现一定的功能. 系统的集成不仅仅能够缩小尺寸，降低成本，更为重要的是能够提高系统的稳定性和可靠性.

由于光学器件和电学器件在结构和工艺上存在许多差别，为了有效地将硅基光学器件和传统 CMOS 电路集成，人们提出了很多种办法. 总体上来说，分为混合集成和单片集成两种.

混合集成是一种最简单的集成方法，它是将光学器件集成在一个硅基底上，电子器件集成在另外一个基底上，然后通过倒装焊或者丝焊的方法将两个芯片连接起来. 混合集成的优点是光学器件和电学器件可以单独优化设计，无需考虑它们互相之间的影响. 它的缺点是无法进行后续的标准封装，因而限制了这种方法的商业应用.

单片集成是用相同的工艺将光学器件和电子器件集成在同一个硅基底上. 由于要同时考虑光学器件和电子器件的设计和工艺，单片集成存在许多困难. 目前，复杂的大规模单片集成还没有实现. 不过，人们提出了许多种单片集成的方案来解决光学器件和电子器件在集成过程中的兼容问题. 根据光学器件在 CMOS 工艺前还是在 CMOS 工艺后，单片集成方案可以分为前端工艺和后端工艺两种. 单片集成是真正大规模实用化的光电集成方案，在下一小节我们将对两种单片集成方案进行较详细的阐述.

9.5.2 前端-后端集成工艺

1. 前端工艺

前端单片集成工艺是指光学器件在电子器件 CMOS 工艺开始的同时集成在芯片上，因此这种单片集成方案中的光学器件和电子器件是混合在一起的. 到目前为止，前端单片集成方案的可行性已经被实验所证明.

绝缘体上硅(SOI)材料是一种广泛应用于集成电路的材料. 它可以实现集成电路中元器件的介质隔离，消除了体硅 CMOS 电路的寄生闩锁效应；采用这

种材料制成的集成电路具有抗辐射性能好、寄生电容小、集成密度高、速度快、工艺简单、短沟道效应小等优点. 同时,绝缘体上硅材料优异的光学性能以及与CMOS集成电路工艺的兼容性,使得低成本的光电子集成成为可能. 由于这一点非常被看好,因而绝缘体上硅材料得到广泛的研究和应用,成为光电子集成的首选材料.

对于大规模集成工艺而言,深紫外光刻能在保证精度的前提下保持高的产量. 为保证集成光波导器件的加工精度,深紫外光刻的波长一般情况下应小于波导中光波长的1/4. 对于SOI单模波导,其有效折射率通常介于2.5～3.0之间. 因此,对于1.55 μm的通信波长,最小的工艺波长为

$$\frac{\lambda}{4n_e}=\frac{1.55\ \mu\text{m}}{4\times 3}=0.129\ \mu\text{m}.$$

因此,0.13 μm的CMOS加工工艺能够满足光电子集成的需要.

电子器件和集成光学器件的兼容性首先体现在SOI材料的顶层硅厚度的选择上. 对于0.13 μm工艺的CMOS晶体管,其要求SOI的顶层硅厚度约为50 nm. 而对于通信波长1550 nm下的光波导器件,单模条件要求SOI的顶层硅厚度约为250 nm. 因此,若采用50 nm顶硅厚度的SOI材料来制备集成光波导器件,光场的限制性将会非常弱. 大部分的光场都不在硅波导中,使得SOI波导器件的弯曲,损耗等性能大大降低. 更为重要的是,由于光场大部分已经不在波导中,实现高效的调制等变得非常困难.

解决这一问题的一种方法是采用双层SOI材料. 如图9.19所示,在这种材料中,顶层硅被分为两层,最上面的一层作为电子器件的有源层,而总厚度作为光波导器件的有源层,这样就解决了厚度上的兼容问题. 双层SOI材料的结构示意图如下图所示. 这种方法的缺点是获得双层SOI原材料非常困难和复杂.

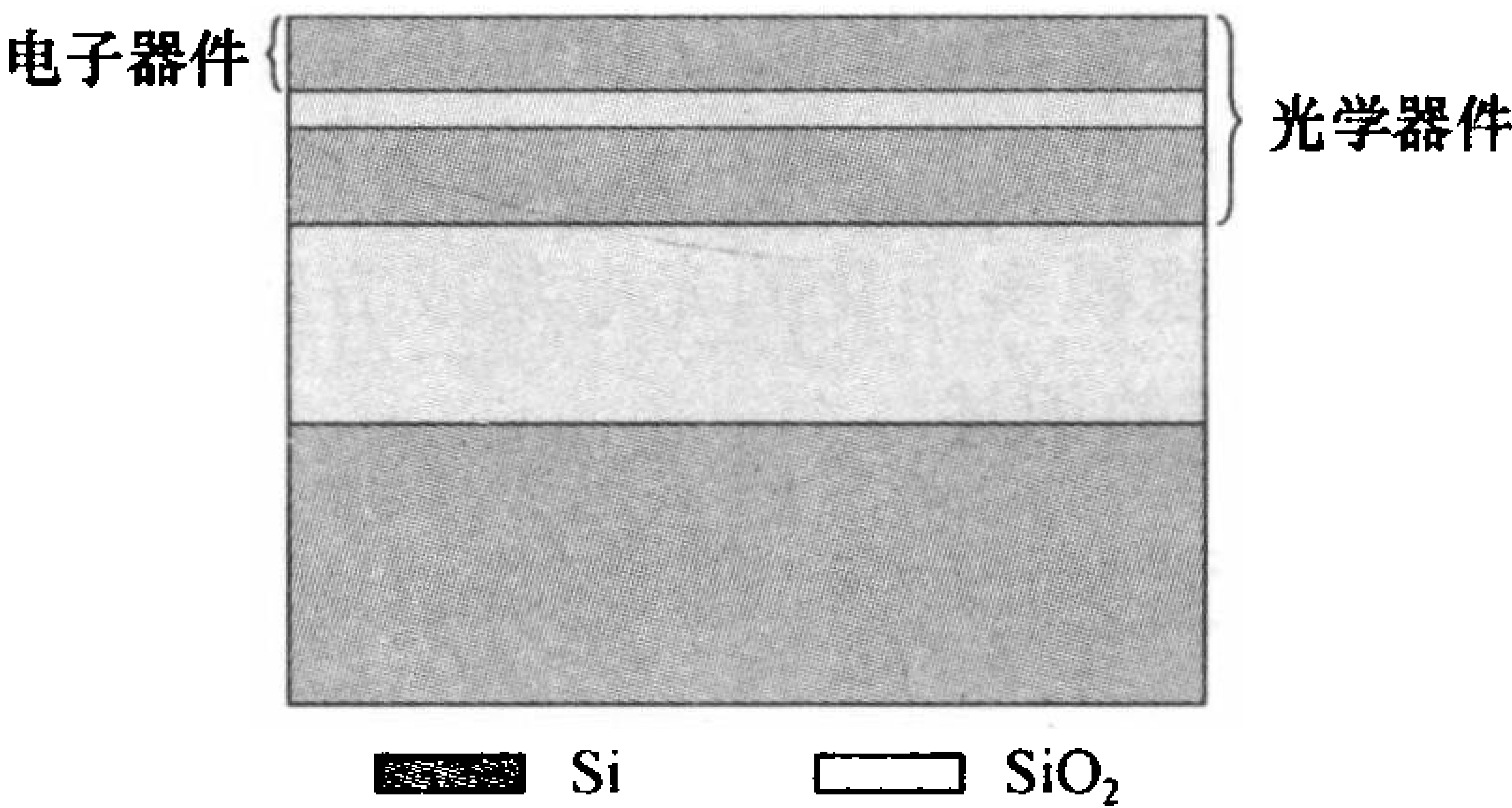


图9.19 双层SOI材料的结构示意图

另外一种方案是采用光学器件要求的SOI材料,电子器件也集成在其中. 此时的晶体管为部分耗尽型,因而其性能会有一定影响,如晶体管的高频特性.

由于光学器件和电子器件在单片集成方案中是集成在一起的,另外还要注

意的光学器件和电子器件之间的相互干扰和影响. 如电学器件中经常要用到掺杂、多晶硅、金属处理等,这些工艺都会影响波导器件的损耗和工作特性. 因此,在设计过程中要多方面考虑结构设计和工艺流程.

2. 后端工艺

光学器件也可以集成在 CMOS 工艺的后端,我们称这种单片集成的方法为后端工艺. 后端工艺通常包含两种方案: ① 在 CMOS 工艺的后端低温沉积光波导器件材料(如非晶硅),并在该材料上制备光学器件;② 在 CMOS 工艺的后端利用三维的集成方法将光学器件集成到基底中. 后端工艺的优点是避免了复杂的前端设计,但它的巨大缺点是改变了已经优化了的后端材料的厚度和性质,需要重新设计和优化.

§9.6 总 结

在本章中,我们主要讨论了硅基光电子器件制备工艺,尤其是常用的、最能影响器件结构和性能的薄膜沉积、光刻和刻蚀工艺. 最后讨论了光电集成存在的问题和可行性方案.

参考文献

[1] 姚汉民,胡松,邢廷文. 光学投影曝光微纳加工技术. 北京: 北京工业大学出版社,2006.

[2] JEOL Ltd., Tokyo, Japan. JBX-9300FS General Statements,2003.

[3] JEOL Ltd., Tokyo, Japan. JBX-9300FS Lithography Operations,2003.

[4] Chang T H P. Proximity effect in electron-beam lithography. J. Vac. Sci. Tech., 1975, 12(6): 1271.

[5] Seo E, Choi B K, Kim O. Determination of proximity effect parameters and the shape bias parameter in electron beam lithography. Microelectronic Engineering, 2000, 53: 305.

[6] Aya S, Kise K, Yabe H, et al. Validity of double and triple gaussian functions for proximity effect correction in X-ray mask writing. Jpn. J. Appl. Phys., 1996, 35: 1929.

[7] Owen G. Methods for proximity effect correction in electron lithography. J. Vac. Sci. Tech. B, 1990, 8(6): 1889.

[8] Parikh M. Self-consistent proximity effect correction technique for resist exposure. J. Vac. Sci. Tech., 1978, 15(3): 921.

[9] Owen G, Rissman P. Proximity effect correction for electron beam lithography by equalization of background dose. J. Appl. Phys., 1983, 54(6): 3573.

[10] Nishida T, Notomi M, Iga R, et al. Quantum wire fabrication by e-beam elithography using high-resolution and high-sensitivity e-beam resist ZEP-520. Jpn. J. Appl. Phys., 1992, 31: 4508.

第十章　硅基光电子学的应用

集成电路彻底改变了现代人的生活，并且随着时代的发展，对于集成电路在集成度、速度、功能及功耗上提出了越来越高的要求. 但是，目前集成电路的发展已经出现了无法逾越的瓶颈，集成光电子学成为目前最有可能取代其的关键技术.

硅基光电子学是研究光电子在以硅为基础的集成器件中的优秀光电特性. 同时由于硅材料具有以下优点：作为高折射率材料可以将光场限制在微纳尺寸；加工工艺与现有 CMOS 工艺兼容；硅元素储量丰富，价格便宜等. 因此，硅基光电子学在实现集成应用方面具有巨大的优势.

前几章论述了硅基光电子材料及基本器件的特性，硅基光电子学的应用覆盖光互连、光通信、光传感和光能源多个领域. 本章将介绍硅基光电子学在上述四个领域的前沿应用实例.

§10.1　光互连——集成高速硅基互连芯片

根据国际半导体技术蓝图（ITRS）发布的未来半导体工艺技术预测，2016 年世界集成电路主流工艺线宽将为 22 nm；2022 年将达到 10 nm. 届时，以硅为基础的微电子技术发展所遵循的摩尔定律将不再适用，器件尺寸和速度逐步趋于物理极限；另一方面，随着处理器速度的提高，电互连所固有的局限性也逐渐显示出来了，如带宽、功耗、时延、串扰、噪声等，已成为集成电路进一步发展的巨大障碍，因此发展电互连的替代或补充连接方式成为必然. 作为下一代互连技术强有力的竞争者，光互连具有宽频带、抗电磁干扰、保密性强、传输损耗低、功耗小等明显优于电互连的特点，是一种极具潜力的电互连替代或补充方案. 同时可以充分利用波分复用（WDM）技术，尤其是密集波分复用（DWDM）技术发挥光互连带宽优势，实现高速、海量信息传输. 本节以集成高速硅基互连芯片为例，讲述硅基光电子学在光互连方面的典型应用.

10.1.1　波分复用原理

WDM 技术是利用单模光纤的带宽以及低损耗的特性，采用多个波长作为载波，允许各载波信道在光纤内同时传输，具体原理如图 10.1 所示.

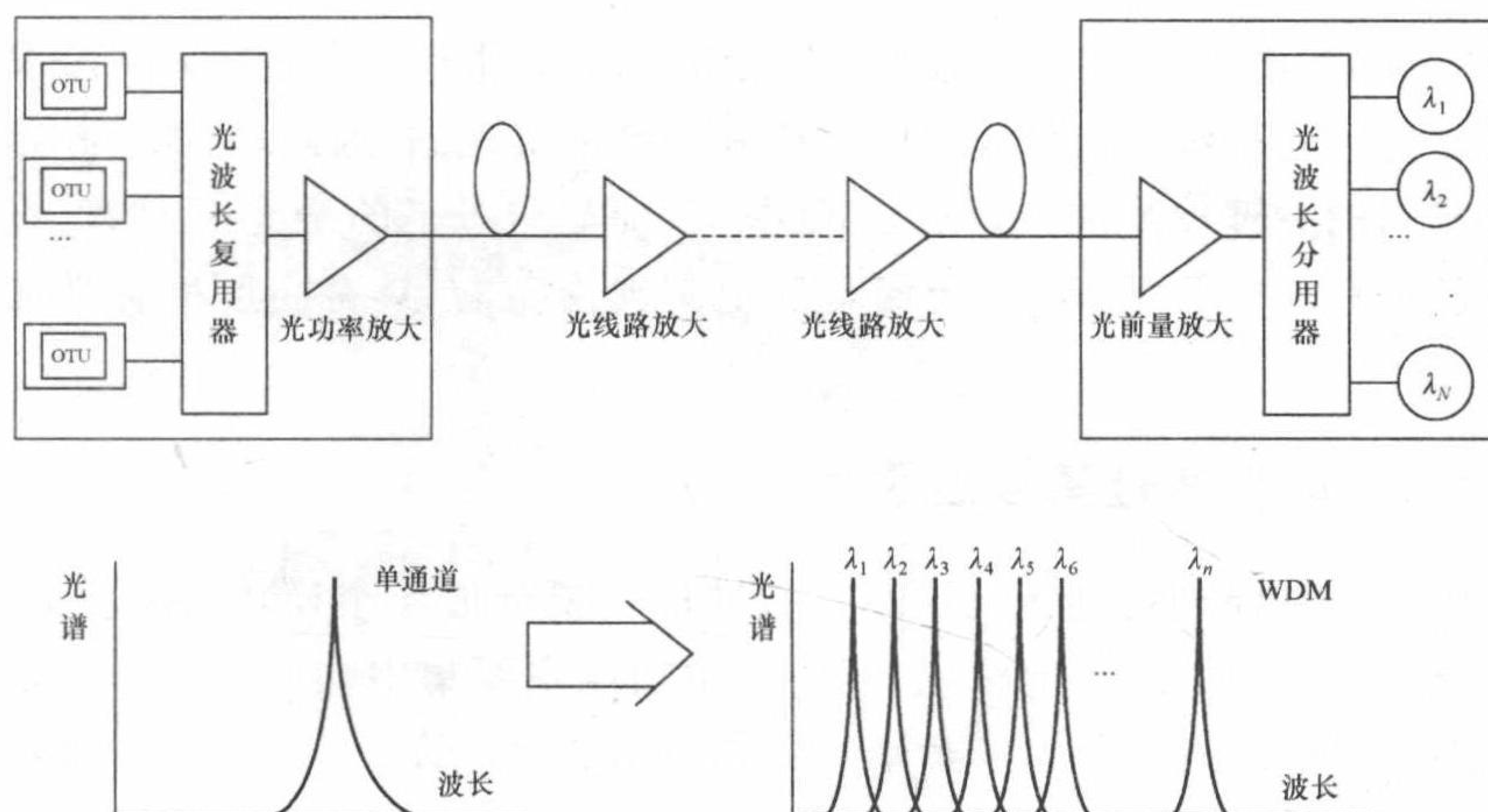


图 10.1 波分复用原理图

WDM 技术具有如下特点：

① 超大容量.

目前使用的普通光纤可传输的带宽是很宽的，但其利用率还很低. 使用 WDM 技术可以使一根光纤的传输容量比单波长传输容量增加几倍、几十倍乃至几百倍. 现在商用最高容量光纤传输系统为 1.6 Tbit/s 的系统，朗讯和北电网络两公司提供的该类产品都采用 1.6 Gbit/s 方案结构. 容量达 3.2 Tbit/s 的实用化系统的开发已具备条件.

② 对数据的“透明”传输.

由于 WDM 系统按光波长的不同进行复用和解复用，而与信号的速率和电调制方式无关，即对数据是“透明”的. 一个 WDM 系统的业务可以承载多种格式的“业务”信号. WDM 系统完成的是透明传输，对于“业务”层信号来说，WDM 系统中的各个光波长通道就像“虚拟”的光纤一样.

③ 系统升级时能最大限度地保护已有投资.

在网络扩充和发展中，无须对光缆线路进行改造，只需更换光发射机和光接收机即可实现，是理想的扩容手段，也是引入宽带业务的方便手段，而且利用增加一个波长即可引入任意想要的新业务或新容量.

④ 高度的组网灵活性、经济性和可靠性.

利用 WDM 技术构成的新型通信网络比用传统的电时分复用技术组成的网络结构要大大简化，而且网络层次分明，各种业务的调度只需调整相应光信号的波长即可实现. 由于网络结构简化、层次分明以及业务调度方便，由此而带来的网络的灵活性、经济性和可靠性是显而易见的.

⑤ 可兼容全光交换.

可以预见,在未来可望实现的全光网络中,各种电信业务的上/下、交叉连接等都是在光上通过对光信号波长的改变和调整来实现的.因此,WDM 技术将是实现全光网的关键技术之一,而且 WDM 系统能与未来的全光网兼容,将来可能会在已经建成的 DWDM 系统的基础上实现透明的、具有高度生存性的全光网络.

10.1.2 集成高速硅基互连芯片

基于 WDM 原理的高速互连芯片按功能主要分成 4 个模块:控制单元、运算单元、光电接口单元、其他单元,其中光学功能主要集中在光电接口模块这一部分,是本小节介绍的重点.芯片示意图如图 10.2 所示.

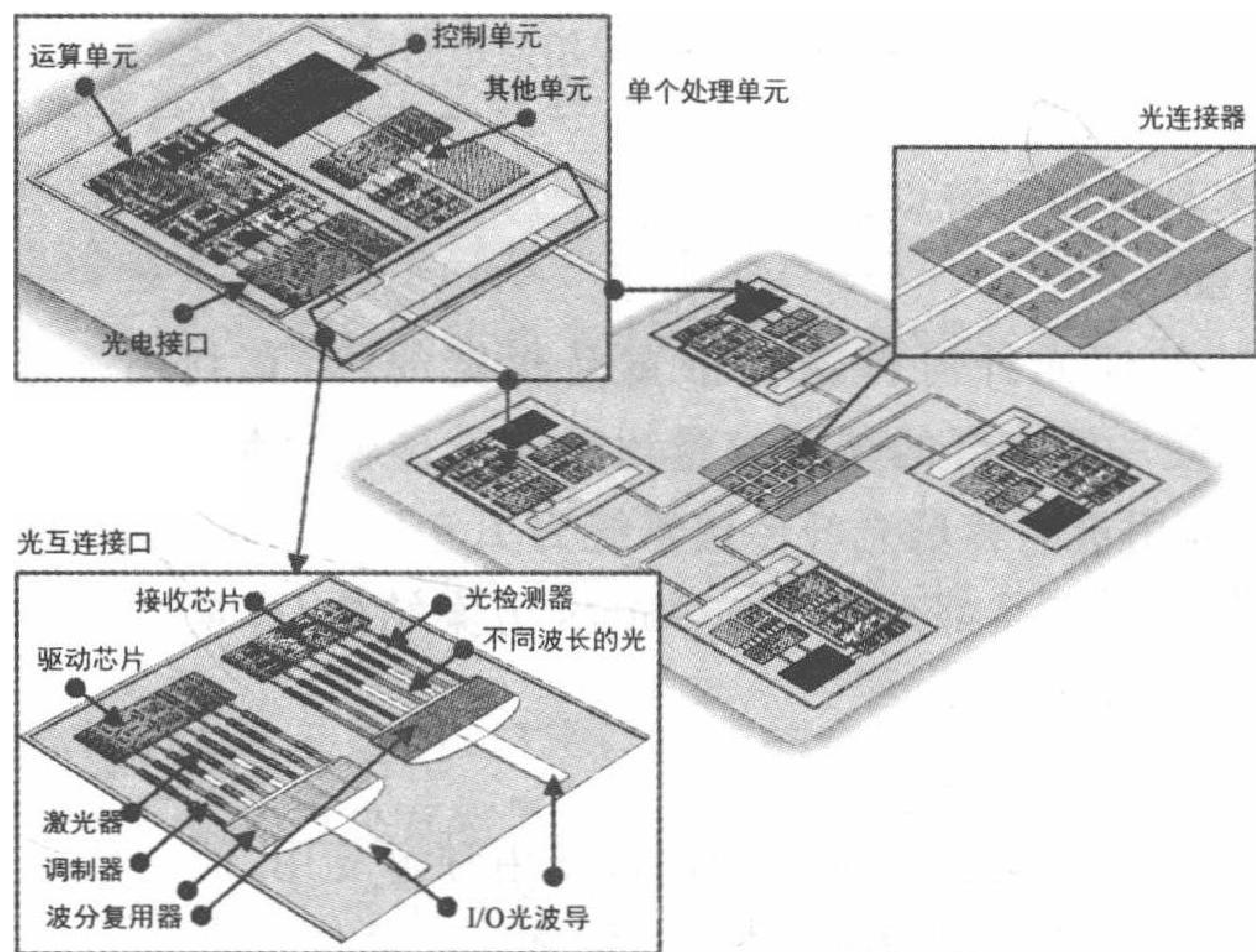
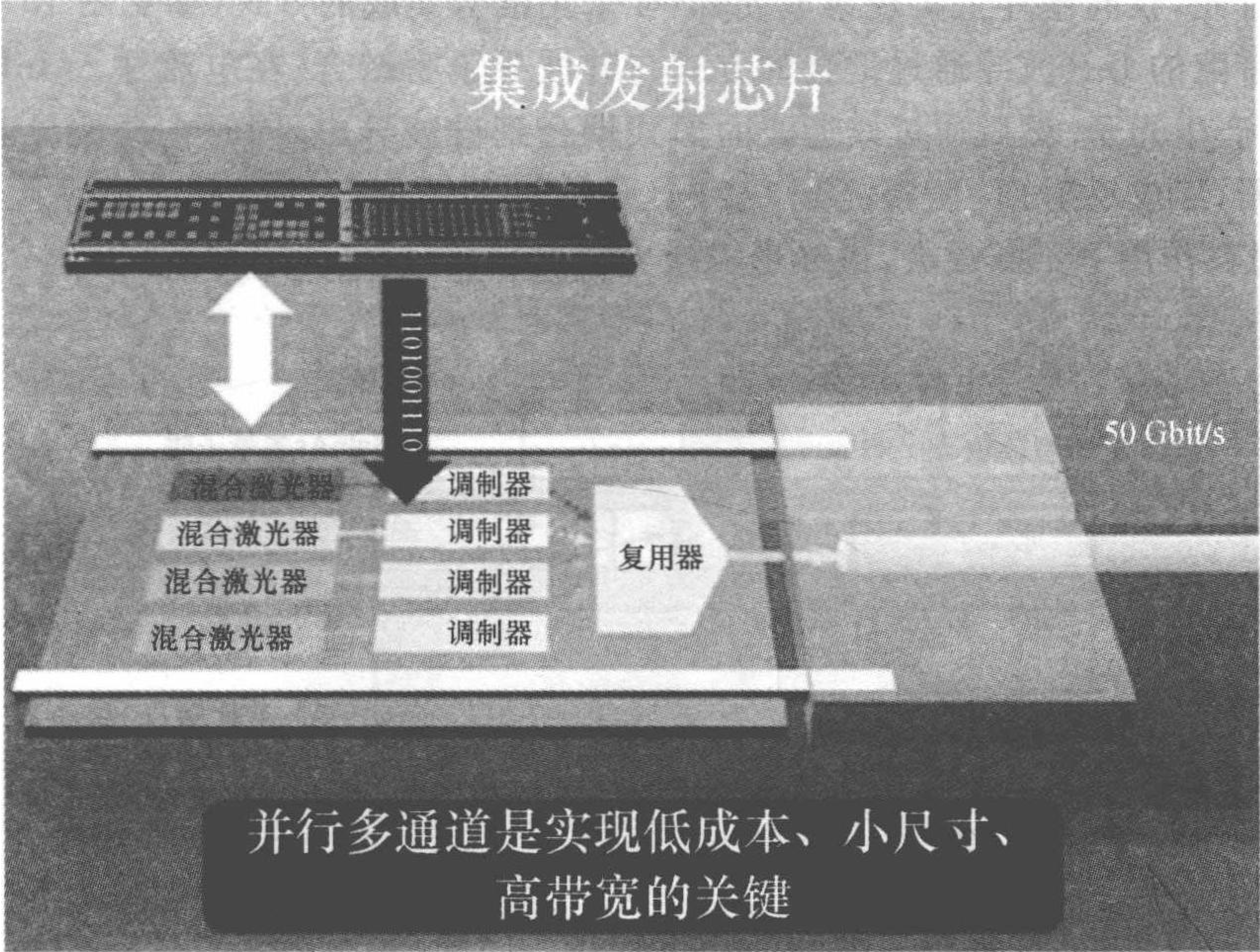


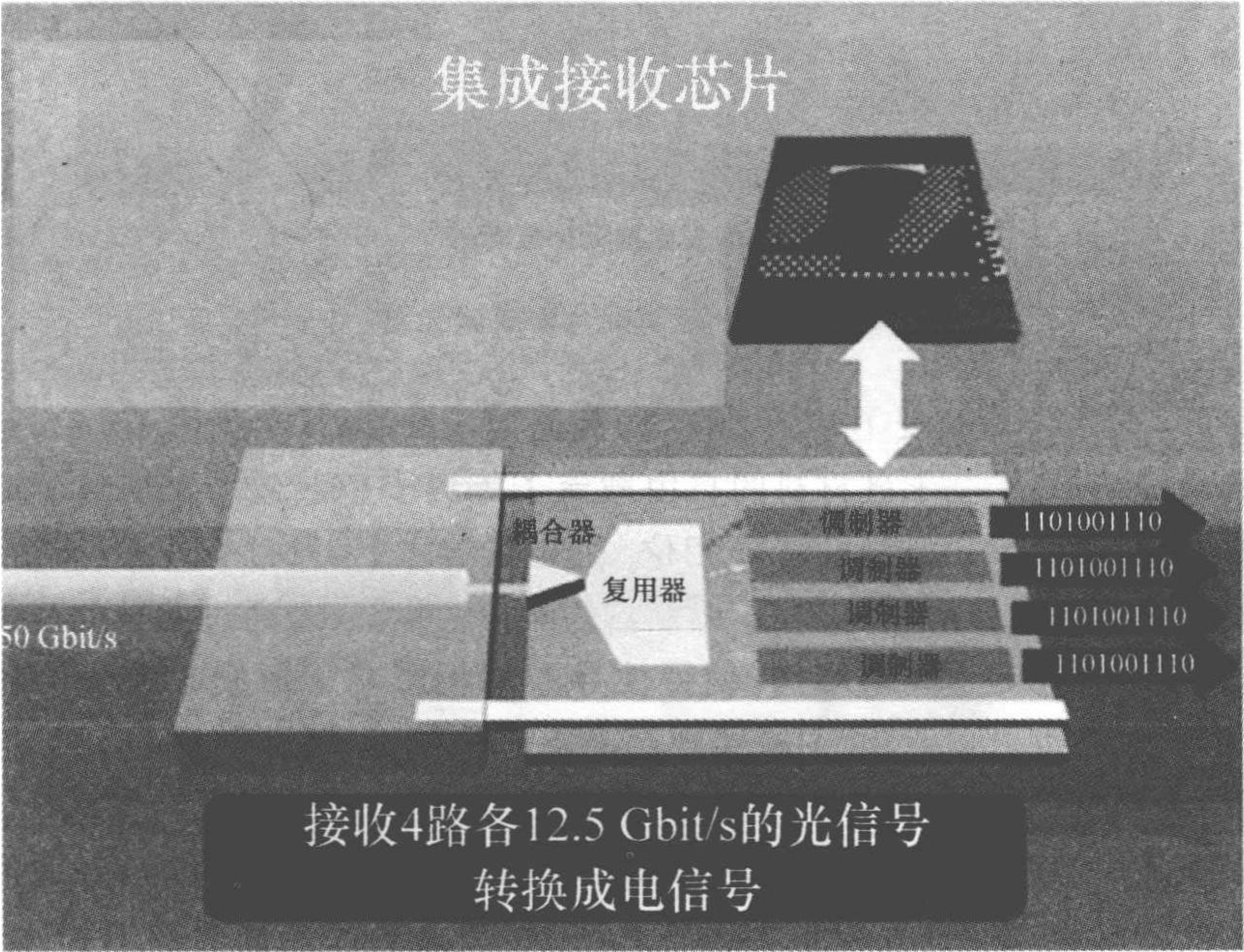
图 10.2 集成高速硅基互连芯片系统结构

根据光学功能区分,光电接口模块有两部分:高速发射和高速接收.高速发射光路先将信号调制到多路光载波上;然后将各路光信号上载到光干线上;高速接收光路将干线信号下载到本地;随后根据波长将多路光载波分离,再进行光电转换后进行信号获取.需要实现上述原理中的功能,具体的光学构成主要有以下单元器件:激光器、调制器、合波/分波器、探测器.

目前实现的集成硅基高速互连芯片来自于 Intel 公司的工作,如图 10.3 所示[1].基于以上单元器件实现 50 Gbit/s 的光互连功能,利用 WDM 原理,可以通过四路复用为例来实现,因此每路的传输速率仅需达到 12.5 Gbit/s,集成芯片重点解决了单元器件的集成及性能问题.



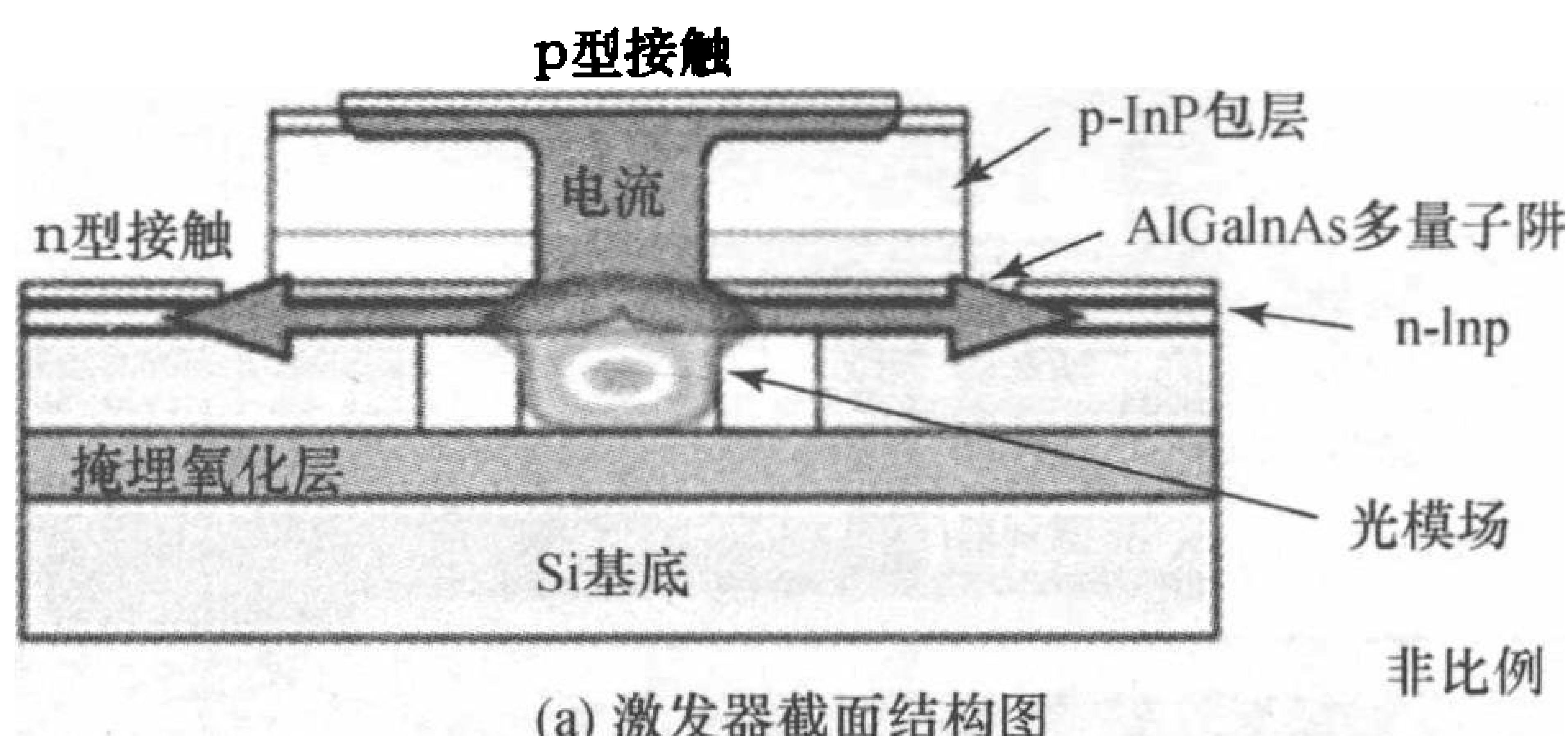
(a)



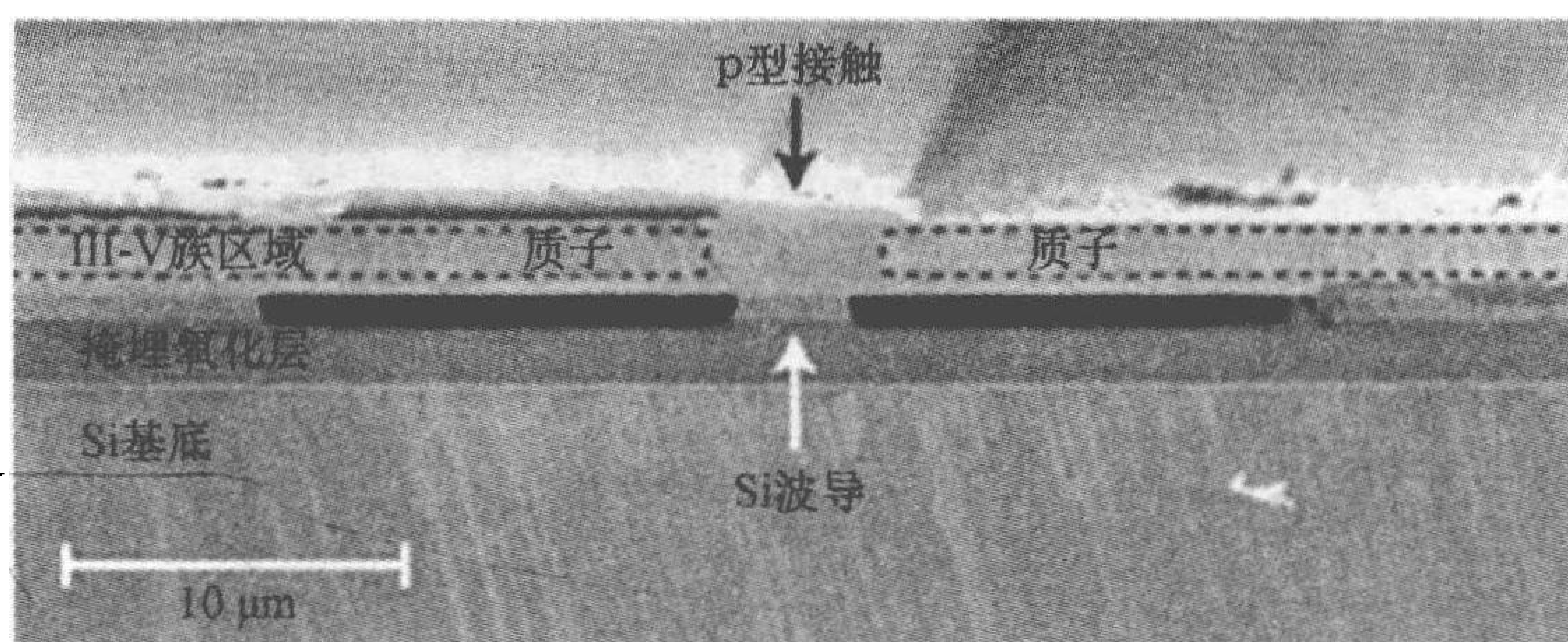
(b)

图 10.3　Intel 50 Gbit/s 硅基光互连芯片

首先，在硅基片上实现集成光源[2]. 采用混合集成的方式（参见图 10.4），将直接带隙的 InP 波导通过键合的方式与硅波导紧密结合. 在电流注入 InP 波导中时，产生光子. 然后通过倏逝波的形式耦合到硅波导中. 同时在硅波导两端刻蚀光栅，形成反射面，使得光场在两个反射面之间来回反射形成谐振腔，并在 InP 材料作用下不断放大，最终形成激光输出.



(a) 激发器截面结构图



(b) 激发光器SEM截面图

图 10.4 硅基混合集成激光器

然后，IBM 公司实现了 40 Gbit/s 的高速硅基调制器（参见图 10.5），是目前报道的最好结果[3]. 基于硅材料的自由载流子等离子色散效应实现对硅波导折射率的调制，从而改变波导的传输相位. 在此基础之上，通过 M-Z 干涉结构，将相位调制转变为强度调制，实现了调制器的调制强度输出. 通过波导的截面图可以看到，基于反向的 p-n 结结构，使得波导在电场影响下形成载流子的耗尽状态，产生较大的折射率变化. 为了实现高的电光调制速率，还特殊设计了行波电极结构，实现阻抗匹配.

集成芯片的高速信号接收，需要硅基集成的光电探测器. 由于硅在通信波段的吸收特性较弱，因此，采用锗来实现光电转化[4]. 集成器件结构如图 10.6 所示，锗材料位于硅波导的上方，通过倏逝场将硅波导中的光场吸收到材料中. 从器件的截面图可以看出，探测器采用的 p-i-n 结构，能提供较高的探测速率，通过这样的器件设计，最终实验测得信号探测速率达到 40 Gbit/s.

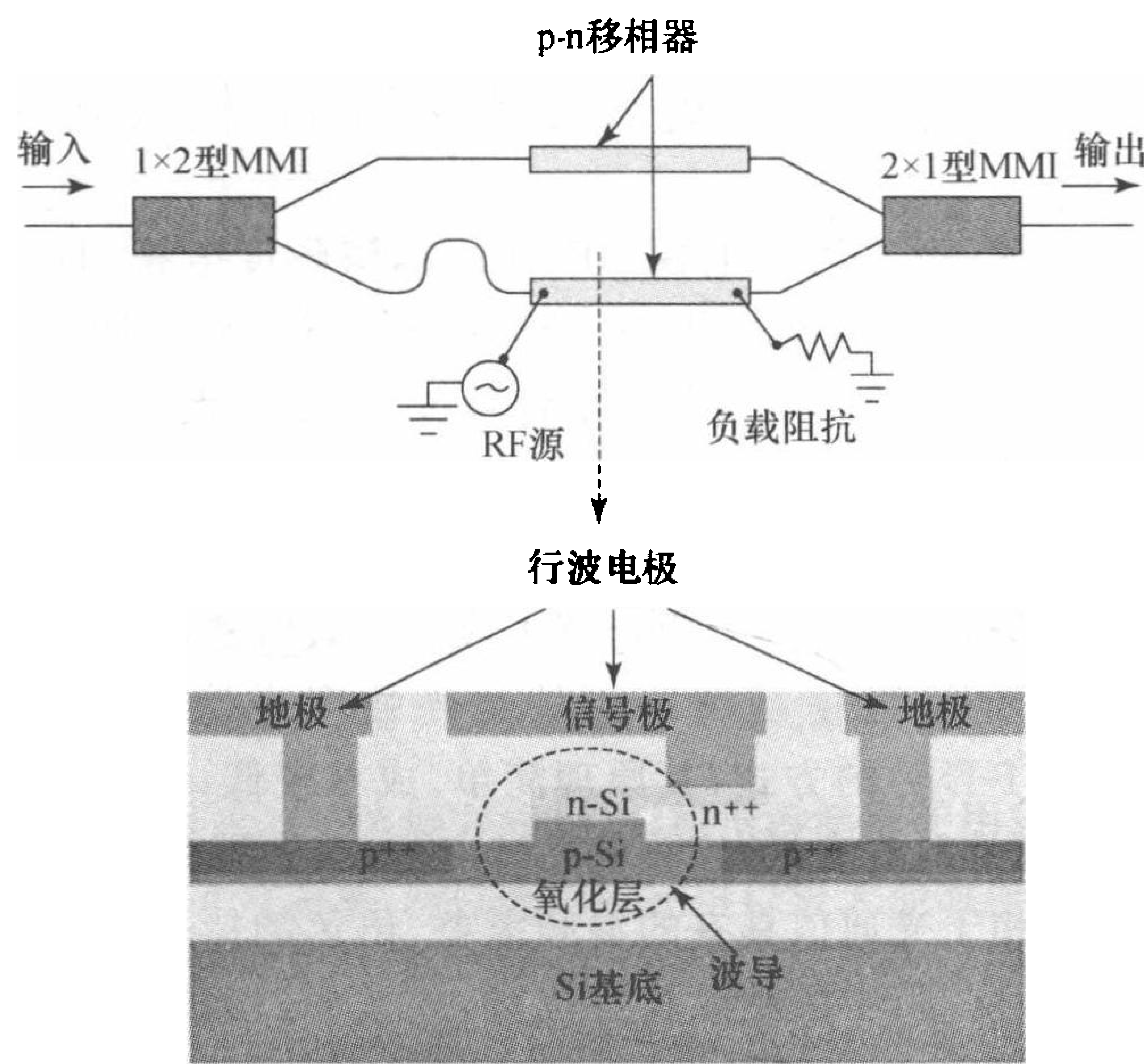
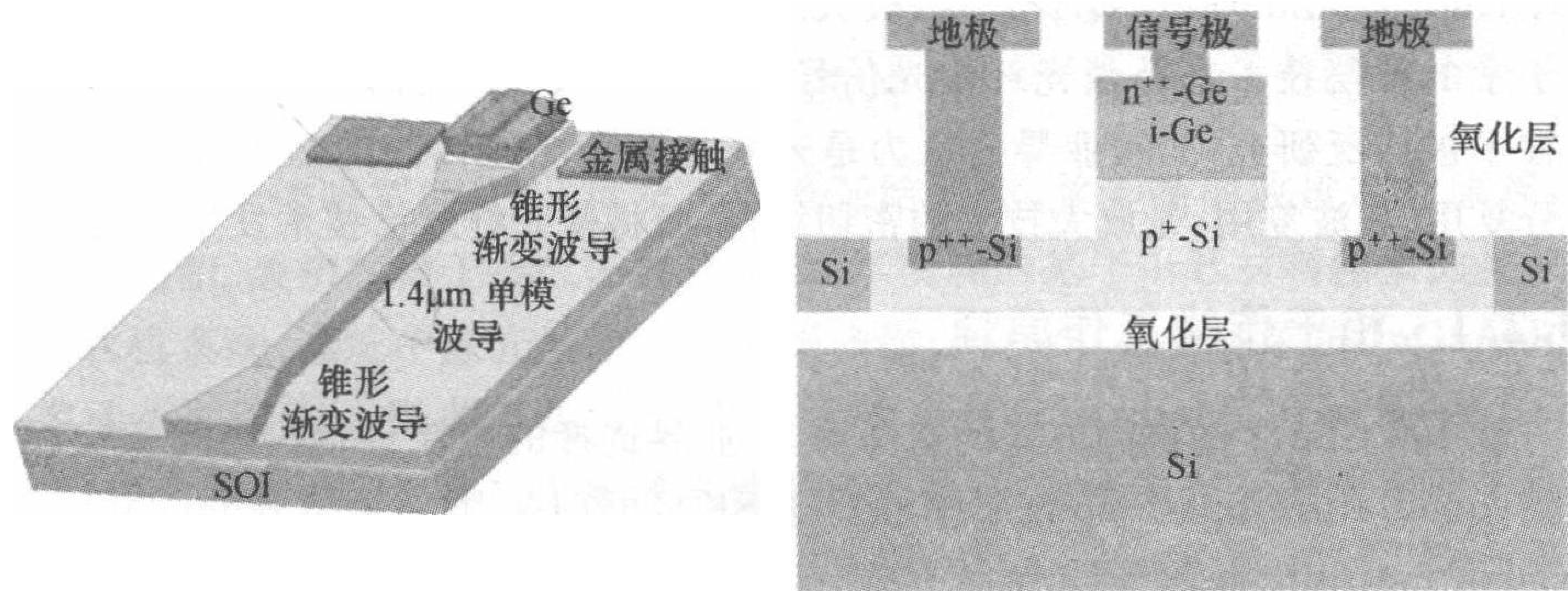
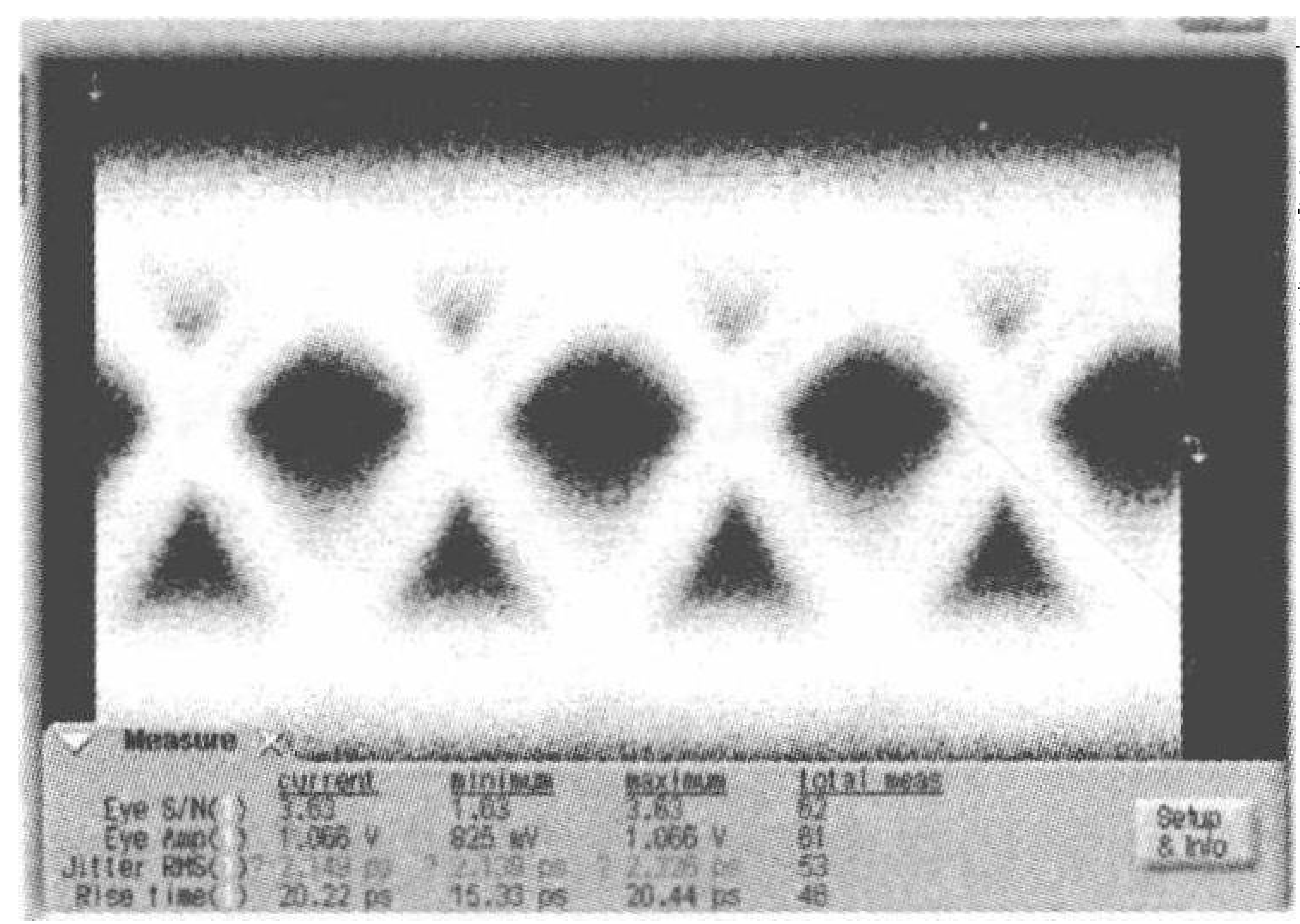


图 10.5　高速集成硅基电光调制器



(a) 波导/探测器耦合结构图　　　　(b) 探测器截面图



(c) 40 Gbit/s光电探测实验眼图

图 10.6　硅基集成高速锗硅探测器

§10.2 光通信——集成 100 Gbit/s 相干传输芯片

过去的三年里面，由于网络电视(IPTV)、高清晰度电视(HDTV)、视频点播(VOD)技术和移动宽带业务的快速发展，使运营商的骨干网络的业务流量持续增长. 到 2012 年，全球的网络业务流量将超过 40 艾可萨字节(exabytes byte, EB)(10^{18}字节). 许多运营商预计网络业务流量年平均增长率达到 50%以上，按此计算，六年后的网络带宽需求将是当前网络的 10 倍以上. 为了应对大容量网络带宽要求，高速率的 100 Gbit/s 传输技术成为光通信领域研究的热点.

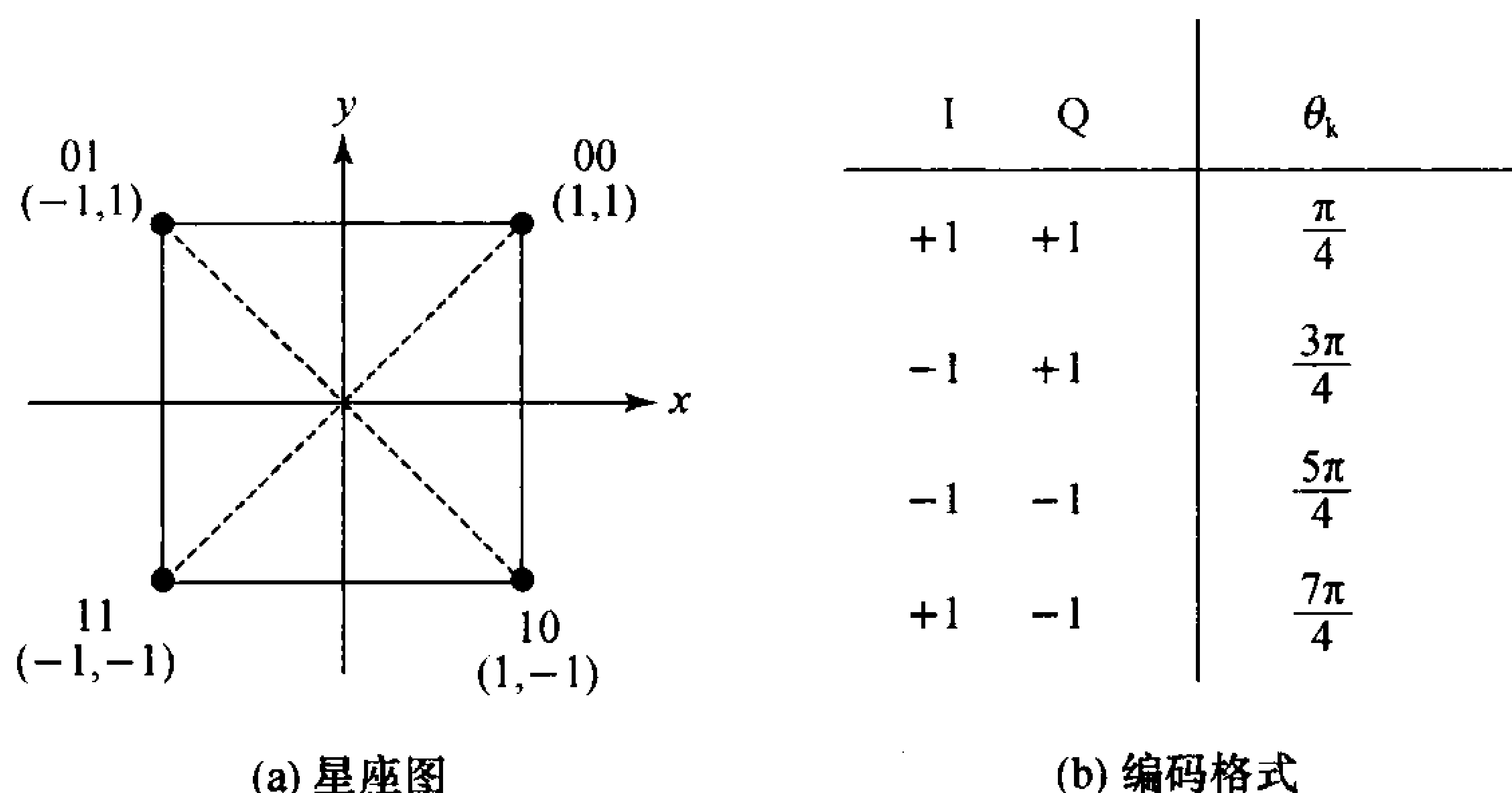
目前实用化的光纤通信系统采用的是光强度直接探测(light intensity direct detection, LT-DD)的方式，其原理简单，成本较低，但是不能发挥光纤通信的优越性，存在频带利用率低，接收机灵敏度差，中继距离短等缺点. 相干光接收技术充分利用了相干光通信具有的混频增益、出色的信道选择性及可调谐的特点，具有接收灵敏度高，中继距离长等优势，并且可以兼容新型高效的调制技术，如相移键控(*n*-phase shift keying, *n*-PSK)、正交幅度调制(quadrature amplitude modulation, QAM)等，极大地提高了传输频谱效率，同时还可以利用电子学的补偿技术来补偿光纤中光信号脉冲的色散效应. 因此，近年来相干光通信技术被广泛研究[1]，被业界公认为是未来光通信高速传输的必备技术，它与光波分复用、载波复用、光放大技术的密切结合，必将使光纤通信技术发生巨大变革.

10.2.1 相干传输工作原理

就 100 Gbit/s 的调制方式的选择而言，业界选择的主流技术仍然是正交相移键控(quadrature phase shift keying, QPSK)，同时采用偏振复用(polarization division multiplexing, PDM)的方式，也就是 PDM-QPSK 的调制方式，该调制方式已经被光纤互联网络论坛(optical interenet working forum, OIF)列为标准.

QPSK 是现代通信中应用最为广泛的数字调制方式之一，如图 10.7 所示. 可以看到，一次同相正交信号(inphase-quadrature, IQ)信号的调制携带了 1/4 的信息，可以表示成 2 bit 的信号. 因此通过 QPSK 信号调制之后，100 Gbit/s 的信号可以通过 50 Gbit/s 的调制来实现. 在通过偏振模复用，调制速率可以进一步减半到 25 Gbit/s.

PDM-QPSK 的信号在接收侧采用相干检测技术可以实现高性能的信号解调. 相干光检测技术，除了检测光信号的幅度，还能检测光信号的相位和偏振态(极化态). 采用相干光技术后，在同样的光谱带宽内就可以传输更大容量的数据. 此外，相干光系统中，光信号的相位信息可用来高效地补偿光传输所引起的信号在电域



(a) 星座图　(b) 编码格式

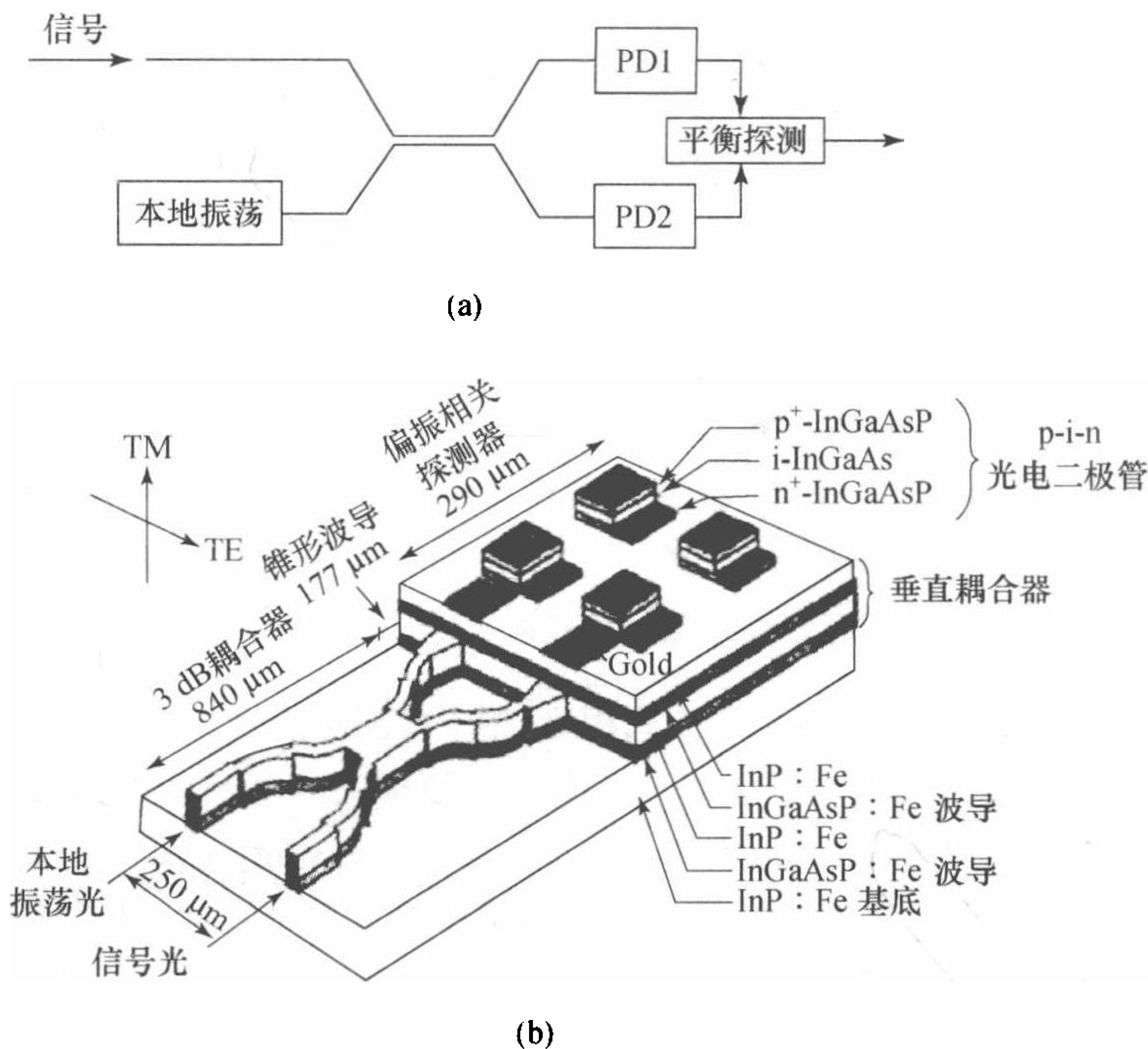
图 10.7　QPSK 原理及编码格式

内的劣化. 与直接解调和差分解调方式相比，相干检测所使用的本地激光器的功率要远大于输入光信号的光功率，所以光信噪比可以极大地改善. 特别是相干检测技术充分利用强大的数字信号处理(digital signal processing, DSP)来处理极化模复用信号，可以通过后续的数字信号处理补偿并进行信号重构，可以还原被传输的信号的特性(极化模、幅度、相位)，大幅度消除光纤带来的传输损伤.

10.2.2　集成 100 Gbit/s 相干接收芯片

集成相干接收芯片从 20 世纪末已经开始研究，并不断的改进芯片原理，提高芯片性能. 最初的相干接收芯片是 1989 年实现的基于 InP 材料体系的集成相干接收机系统，系统中除光波导及光电探测器(PD)外，还分别采用 InGaAsP 材料多量子阱结构[5]和多电极分布式反馈激光器(distributed feed back laser, DFB)激光器结构[6]实现片上集成本地振荡光源，但这些方案都是采用非平衡探测的方式，接收机的灵敏度相对较低. 1990 年，T. L. Koch 和 R. J. Deri 等均采用平衡探测的方案，极大地提高了系统的灵敏度[7,8]. 随后，R. J. Deri 等在前面工作基础上进一步实现了偏振复用的相干接收机，其具体结构如图 10.8 所示，该结构也是基于 InP 材料体系的波导及光探测器实现方案[9]. 首先制作单模波导，将信号光和本地振荡光耦合进器件；然后经过 3 dB 耦合器实现干涉分束；最后 TE 和 TM 的光信号将分别进入两组平衡光电探测器中转换成电信号，完成偏振复用的相干接收.

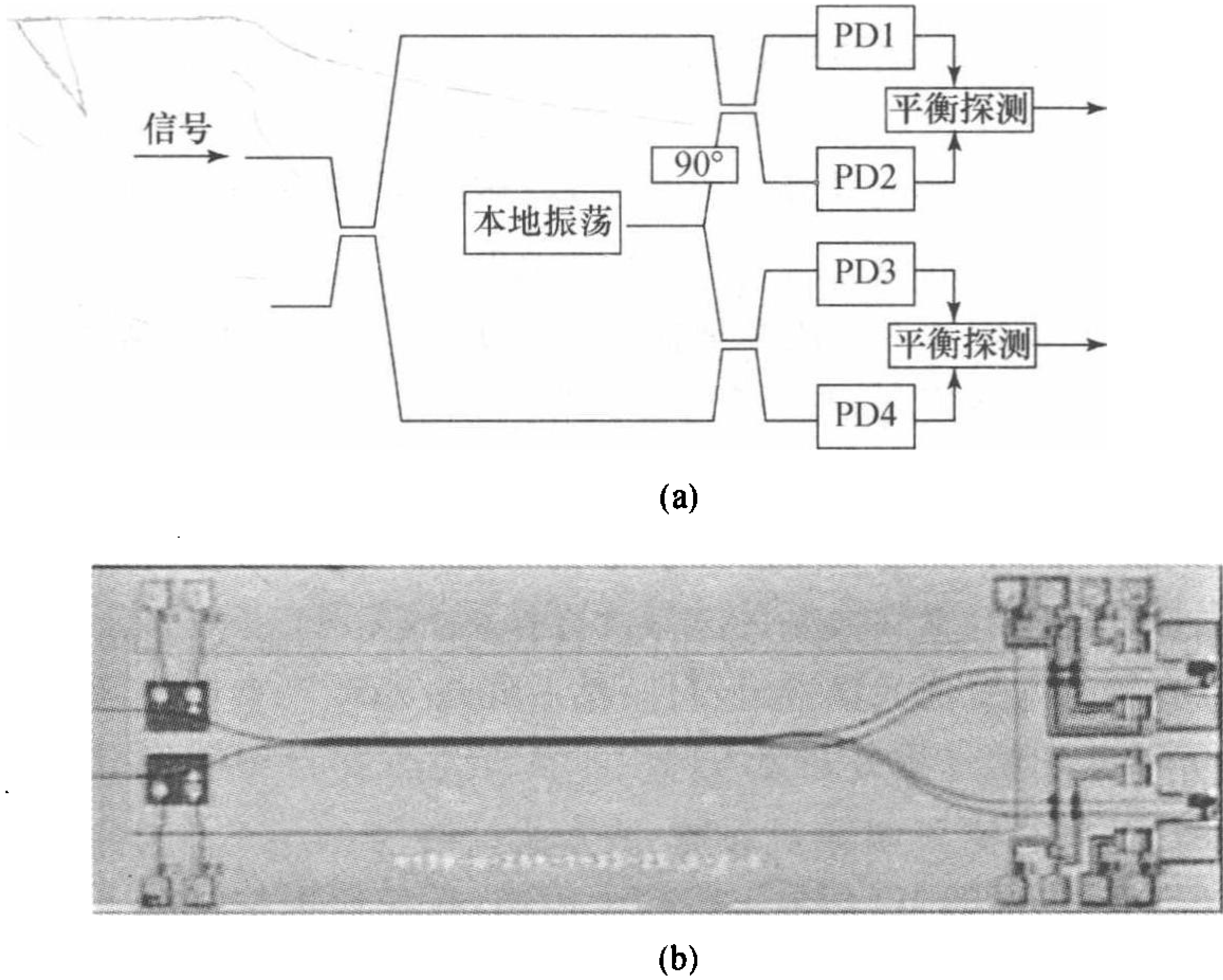
目前我们主要采用的 IQ 信号进行相干传输，在单偏振的情况下，实现的集成芯片如图 10.9 所示[10]. 该集成系统基于 InP 材料体系结构，利用模式转换器将光纤的模式转化为波导的耦合，提高光纤输入的耦合效率. 通过 2×4 型多模干涉仪(MMI)结构实现信号光与本地振荡光的混频，输入如图 10.9(a)所示的四路信号. 最后通过 InP 的 p-i-n 型探测器进行平衡接收.



(a)

(b)

图 10.8 单偏振相干接收系统



(a)

(b)

图 10.9 单偏振 IQ 信号相干接收系统

在偏振复用 IQ 信号相干接收系统中，同时存在两个正交偏振态，需要分别进行相干接收，原理如图 10.10(a)所示. 在单偏振 IQ 信号相干接收系统的基础

上，增加了偏振分束器及一套 IQ 信号相干接收光路. 目前一种实现方案基于 InP 材料体系[11]，其结构图如 10.10(b)所示，偏振控制器(polarization beam splitter, PBS)采用 MZI 结构，混频器(hybrid)采用 2×2 型 MMI 结构，实现对 112 Gbit/s 信号的相干接收. 目前报道的另外一种实现方案是由 Bell 实验室在硅材料上实现的相干接收系统，如图 10.10(c)所示[12]. 通过光栅将光纤中的信号耦合到波导中，并进行偏振分束，通过热电极使其中两个偏振的一路信号产生 90°相移，接着经过 MMI 干涉，利用锗硅的光电探测器的平衡接收将其转化为电信号.

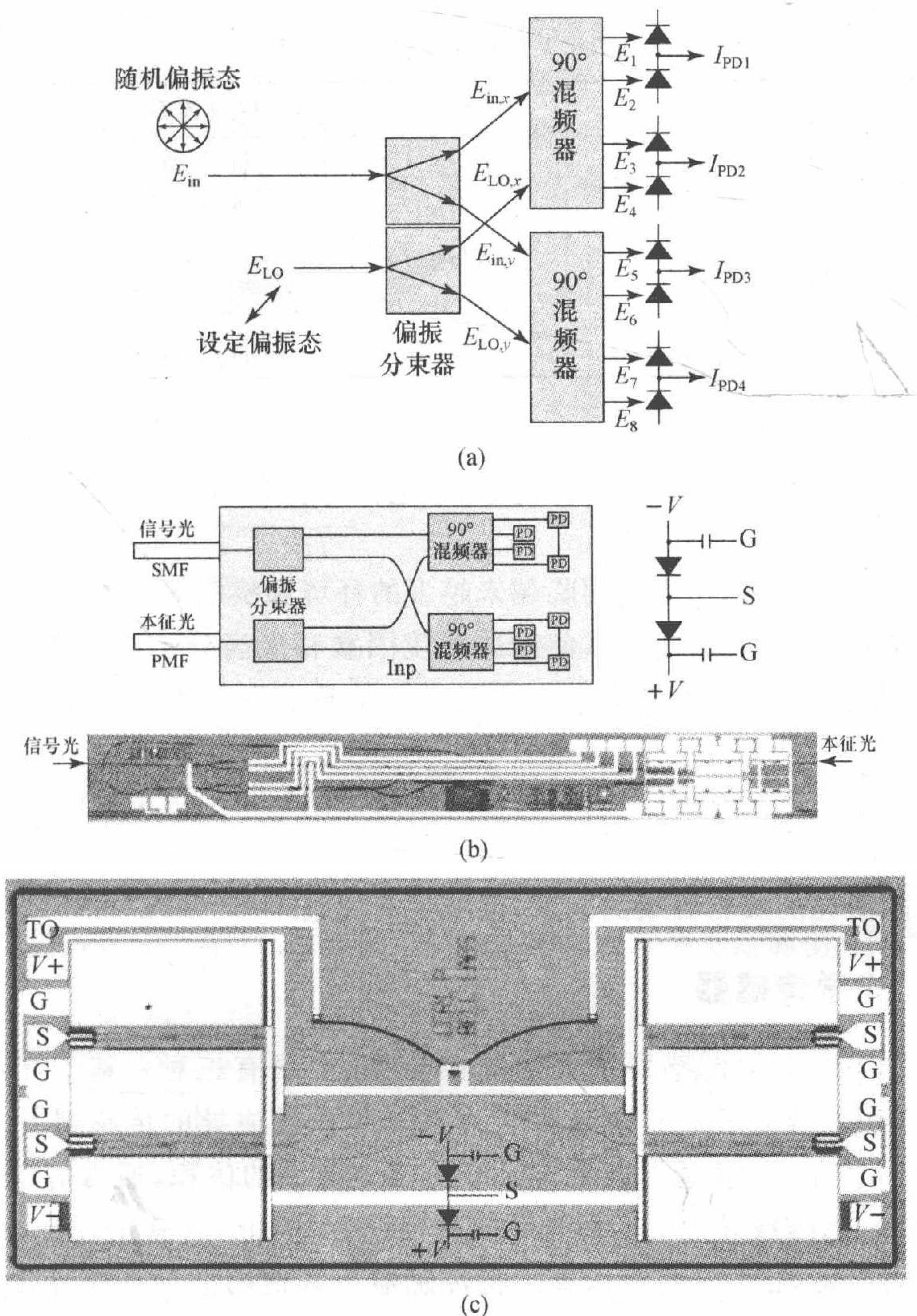


图 10.10　偏振复用 IQ 信号相干接收系统

针对以上几种相干接收芯片的具体实现，从性能、尺寸、材料体系等方面对其进行比较如表 10.1 所示. 可以发现，目前研究的集成相干光接收机系统主要是基于Ⅲ-Ⅴ族材料体系实现的. 该方案可以利用优良的Ⅲ-Ⅴ族材料特性实现片上集成的高性能有源光电子探测器件，然而其加工工艺复杂，器件尺寸较大，且成本昂贵，无法实现大规模集成化. 与Ⅲ-Ⅴ族材料比较而言，硅材料器件尺寸小，相比其他材料成本更低，且硅基光电子集成工艺与目前业界成熟的 CMOS 工艺兼容，具有丰富的工艺积累，这些都使得硅基工艺具有非常大的产业潜力. 因此采用硅基集成技术实现相干接收机系统是实现芯片化、实用化、产业化的最佳方案.

表 10.1 相干接收系统比较

	性能	材料体系	尺寸
非 IQ 信号	2×400 Hz	Ⅲ-Ⅴ族	1.3 mm 长
单偏振 IQ 信号	50 Gbit/s	Ⅲ-Ⅴ族	5.7×1.3 mm^2
偏振复用 IQ 信号	43 Gbit/s 112 Gbit/s	硅基/ Ⅲ-Ⅴ族	3.6×1.6 mm^2
密集波分复用偏振复用 IQ 信号	4×43 Gbit/s	Ⅲ-Ⅴ族	6.0×3.3 mm^2

§10.3 光传感——集成光学传感片上系统芯片

近年来，丰富的物质生活面临越来越多的环境和疾病等方面的健康威胁，人们提出了迫切需求，对影响身体健康的环境因素和疾病分子进行微量检测与监控. 在这一社会需求下，发展面向生命科学的信息获取技术——光学传感技术，成为研究热点.

光学传感技术具有响应快速(可实现实时监测)、抗电磁干扰等优点. 期望通过光学传感技术的研究，能够对有害物质和疾病分子进行准确检测与监控，及早控制并消除潜在的威胁健康的因素，保障生活健康.

10.3.1 光学传感器

目前发展的光学传感器按光学结构分类，主要有三种：基于光纤的传感器[13]、基于表面等离子体的传感器[14]和基于集成光波导的传感器[15]. 其中，基于集成光波导的光学传感器具有比光纤传感器更小的体积，通常的单模波导宽度为 0.5 μm(以硅波导为例)，而光纤的直径约为 10 μm；具有比表面等离子体传感器更高的可集成度，并且可以直接将传感器与其他的波导功能器件集成到同一芯片上[16]. 同时，由于光波导的可集成特性，将其与微流控技术相结合，在实现下一代片上智能传感系统方面具有巨大的潜力[17].

目前基于光学波导的传感功能器件有阵列波导光栅(AWG)[18]、MMI[19]、耦合器(coupler)[20]、法布里-珀罗(F-P)干涉仪[21]等,其中传感灵敏度较高的主要有三种:光子晶体[22]、MZI[23]、环形谐振腔[24].这三种结构相比较,光子晶体加工复杂度较高,MZI的尺寸较大,环形谐振腔具有尺寸小的优势[25],并且由于其高品质谐振特性[26],能够实现可集成的高灵敏传感器.

10.3.2　微纳环形谐振腔光学传感芯片

根据光波导基本传感原理,微纳环形谐振腔能够形成高灵敏度的传感芯片.在有被测物质加载的情况下,波导的有效折射率会发生改变[27].从而导致谐振特性发生改变,具体表现为谐振谱线的漂移,如图10.11所示[28].在通常的传感过程中,谐振谱线发生漂移,根据检测物理量的不同,有两种基本传感方法,一种是基于谐振波长变化的传感,另一种是基于输出强度变化的传感[29].

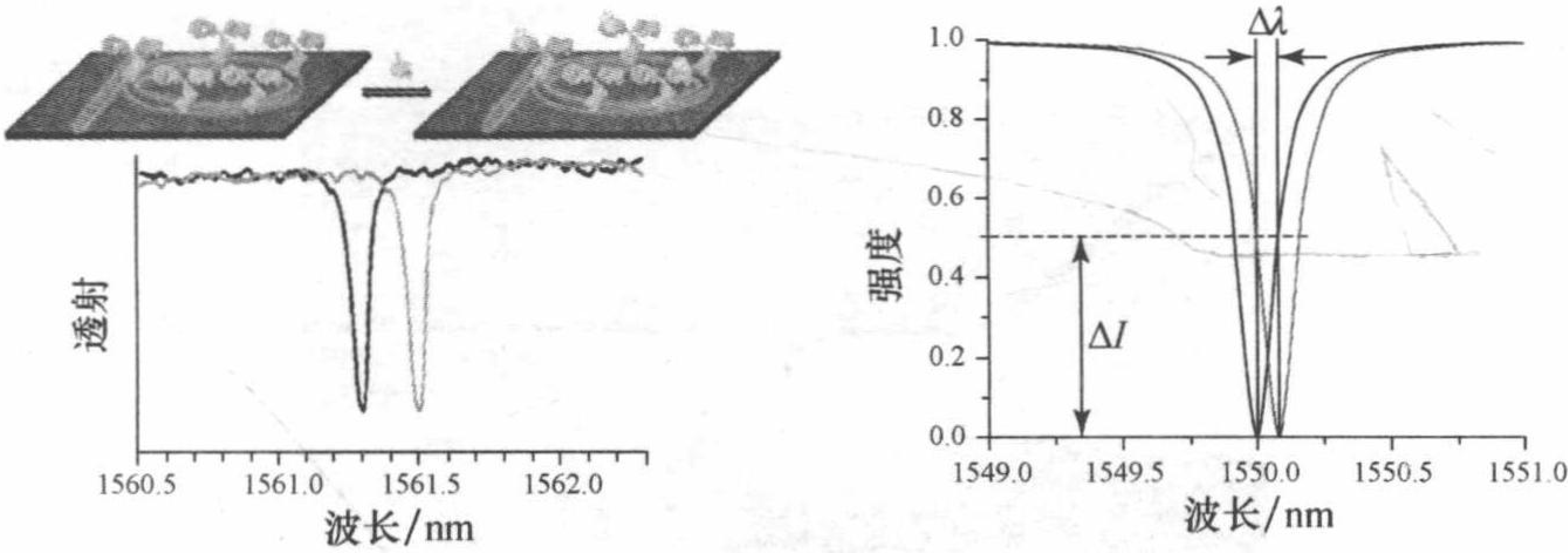


图10.11　基本传感原理

根据微环谐振腔的基本原理,通过探测被测物的微小的折射率改变,从而进行传感.从被测物折射率的改变方式分类,主要有三种:均匀体物质传感、层状物质传感和颗粒物质传感,如图10.12所示.均匀体物质传感的对象通常为不同种类、不同浓度的溶液[30];层状物质传感的对象通常为层状分布的被测物分子,例如DNA分子、蛋白质分子等[31];颗粒物质传感的对象通常为特定的物质颗粒[32].

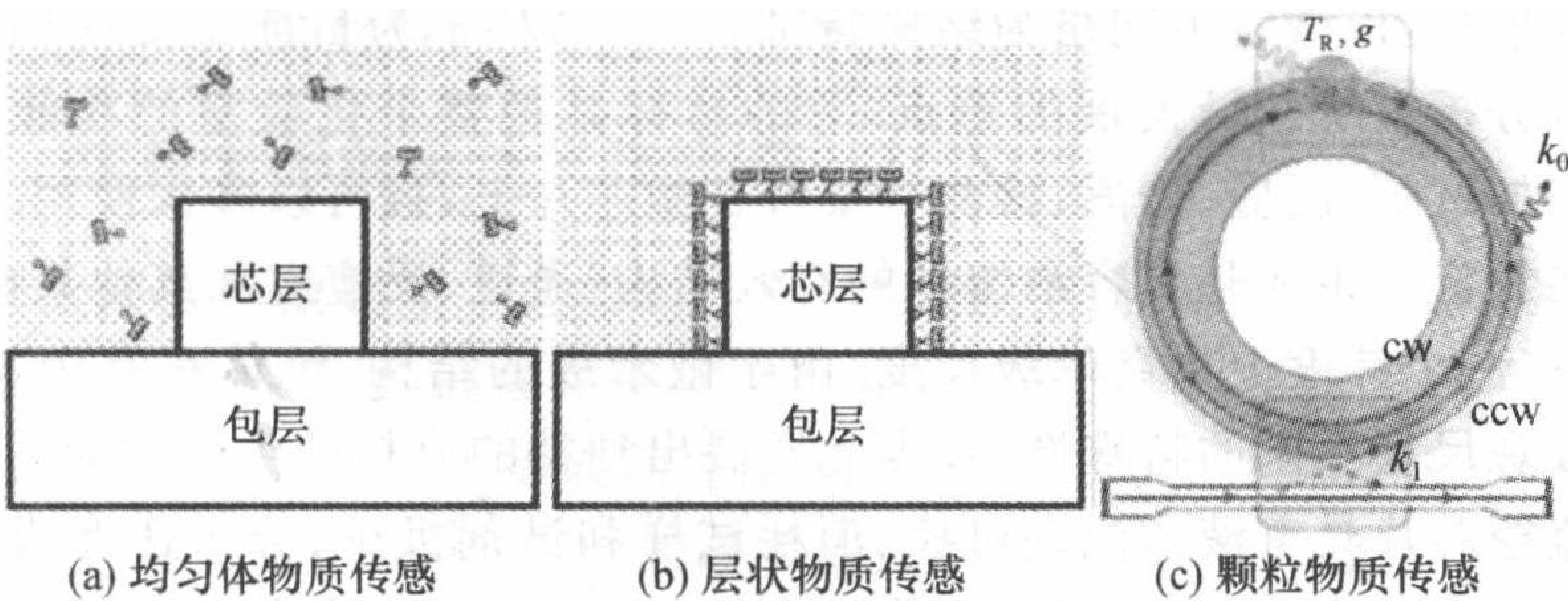


图10.12　传感功能

均匀体物质传感中被测溶液的浓度、种类的改变，导致波导上包层的折射率的改变，从而影响在环形波导中传输的光场的有效折射率；层状物质传感中被测物分子的种类及浓度的不同，导致层状物质的厚度及折射率不同，从而导致环形波导有效折射率发生改变. 这两种传感方式均通过折射率的改变使谐振谱线发生漂移. 颗粒物质传感的原理略有不同，在传感中被测颗粒的大小及数量的不同导致其对光场地散射效果不一样，使得谐振谱线的形状和位置均发生改变，因此实现传感[32].

集成片上光学传感系统芯片的实现，必须将光学传感器件与微流控通道（参见图 10.13）结合起来. 微流（microfluidic）控芯片技术是把生物、化学、医学分析过程的样品制备、反应、分离、检测等基本操作单元集成到一块微米尺度的芯片上，自动完成分析全过程. 由于它在生物、化学、医学等领域的巨大潜力，已经发展成为一个生物、化学、医学、流体、电子、材料、机械等学科交叉的崭新研究领域.

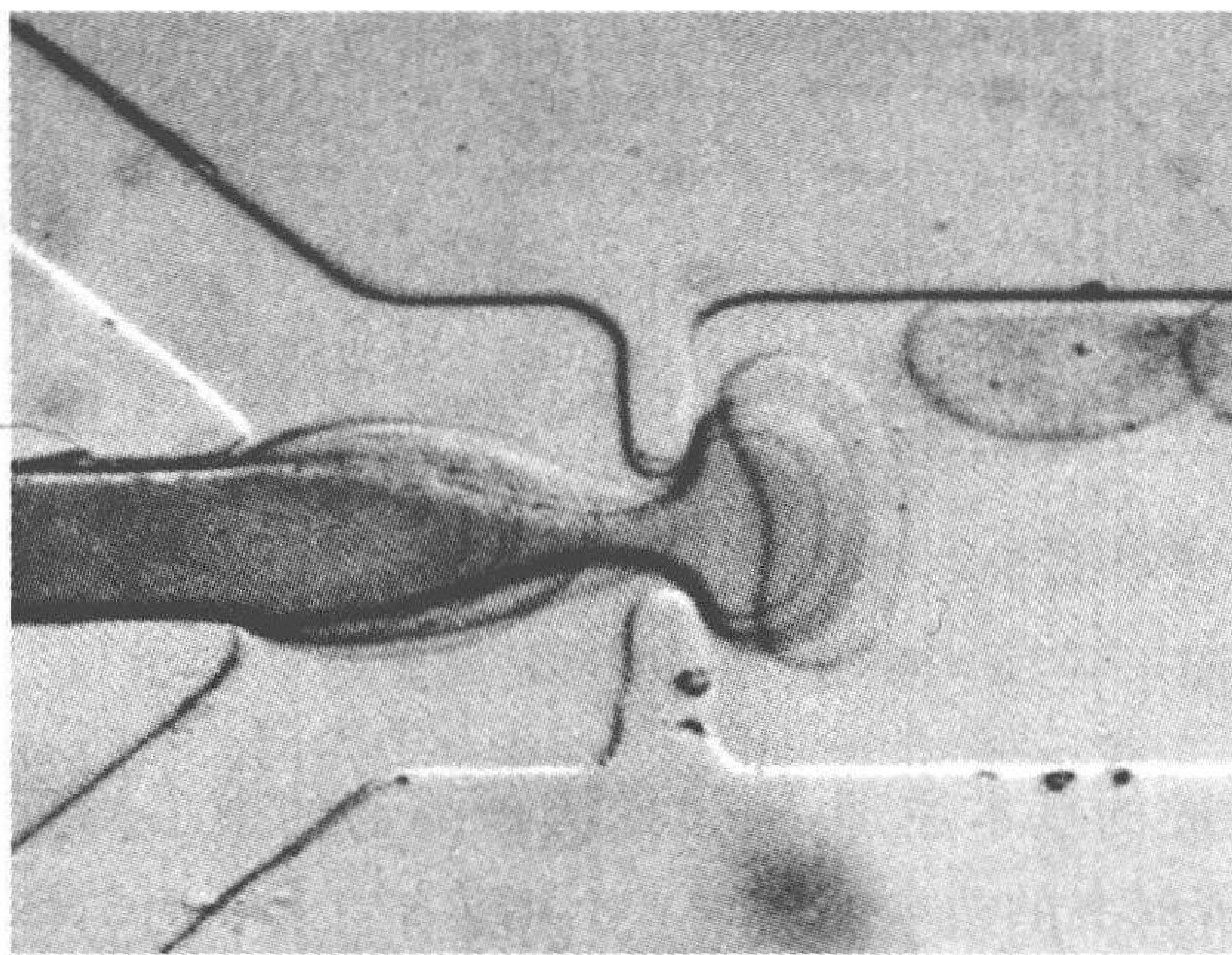


图 10.13 微流控通道

微流控芯片分析以芯片为操作平台，同时以分析化学为基础，以微机电加工技术为依托，以微管道网络为结构特征，以生命科学为目前主要应用对象，是当前微全分析系统领域发展的重点. 它的目标是把整个化验室的功能，包括采样、稀释、加试剂、反应、分离、检测等集成在微芯片上，且可以多次使用.

其装置特征主要是其容纳流体的有效结构（通道、反应室和其他某些功能部件）至少在一个纬度上为微米级尺度. 由于微米级的结构，流体在其中显示和产生了与宏观尺度不同的特殊性能. 因此发展出独特的分析产生的性能.

微流控芯片具有液体流动可控、消耗试样和试剂极少、分析速度成十倍，甚至上百倍提高等特点，它可以在几分钟，甚至更短的时间内，同时分析上百个样品，并且可以在线实现样品的预处理及分析全过程.

利用这样的技术制作的微纳环形谐振腔光学传感芯片，如图 10.14 所示[16,17]. 通过微通道分别控制流经各个不同环形谐振腔器件的被测物，控制被测物与介质波导结合后，测量光学性能的变化，从而得到传感信息. 多个微纳环形谐振腔的阵列可以实现多种物质的同时测量，并且通过与参考环的比较，能够排除温度对器件的影响.

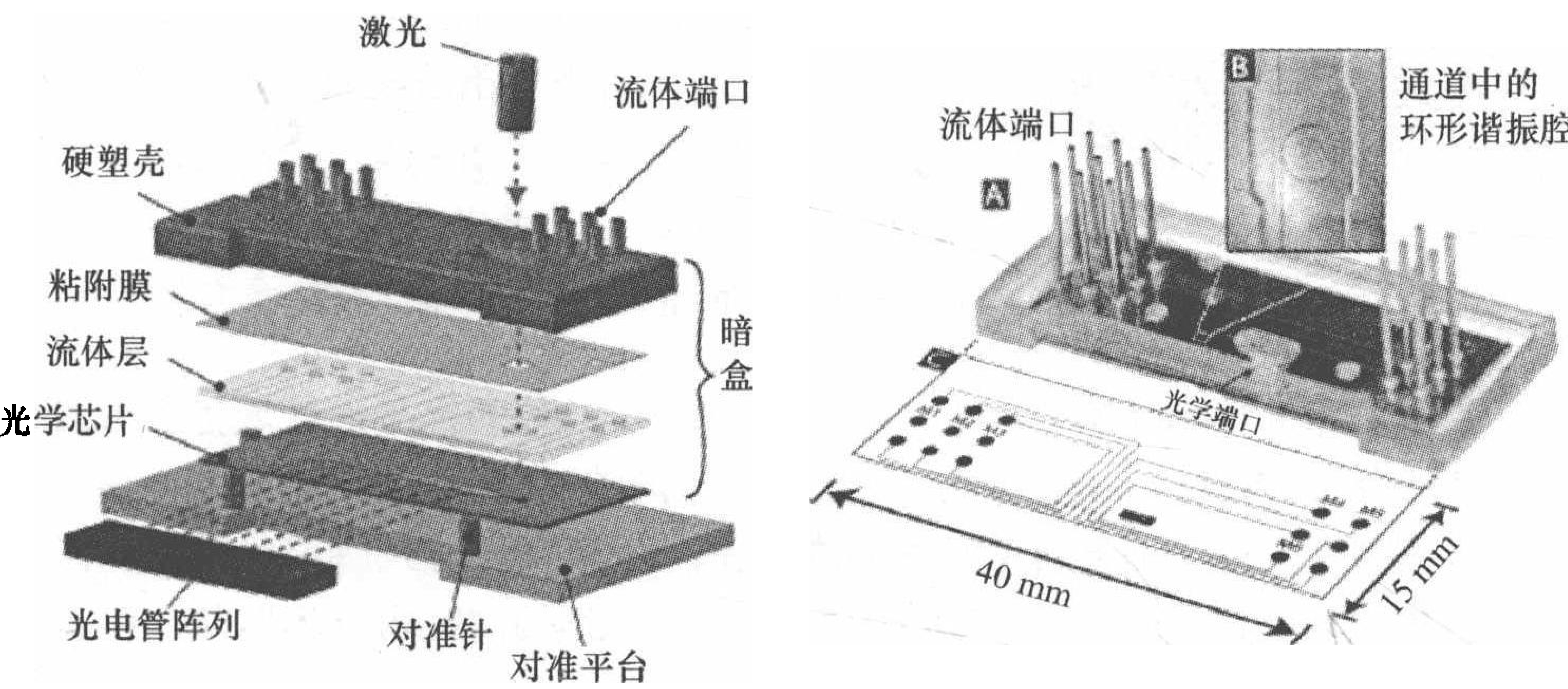


图 10.14　微纳环形谐振腔光学传感芯片[16,17]

§ 10.4　太阳能电池

太阳能是人类取之不尽用之不竭的可再生能源. 也是清洁能源，不产生任何的环境污染. 在太阳能的有效利用当中；太阳能光电利用是近些年来发展最快，最具活力的研究领域，是其中最受瞩目的项目之一. 为此，人们研制和开发了太阳能电池. 对太阳能电池材料一般的要求有：① 半导体材料的禁带不能太宽；② 较高的光电转换效率；③ 材料本身对环境不造成污染；④ 材料便于工业化生产且材料性能稳定. 基于以上几个方面考虑，硅是最理想的太阳能电池材料，这也是太阳能电池以硅材料为主的主要原因.

10.4.1　硅基太阳能电池工作原理

硅基太阳能电池利用了硅材料的光伏原理. 当硅晶体中掺入其他的杂质，如硼、磷等，当掺入硼时，硅晶体中就会存在着一个空穴，它的形成可以参照图 10.15. 图 10.15(a)中，正电荷表示硅原子，负电荷表示围绕在硅原子旁边的四个电子. 如图 10.15(b)所示，灰色的表示掺入的硼原子，因为硼原子周围只有 3 个电子，所以就会产生如图 10.15(b)所示的深色的空穴，这个空穴因为没有电子

而变得很不稳定，容易吸收电子而中和，形成 p 型半导体. 同样，如图 10.15(c)所示掺入磷原子以后，因为磷原子有五个电子，所以就会有一个电子变得非常活跃，形成 n 型半导体，图 10.15(c)中浅色的为磷原子核，深色的为多余的电子.

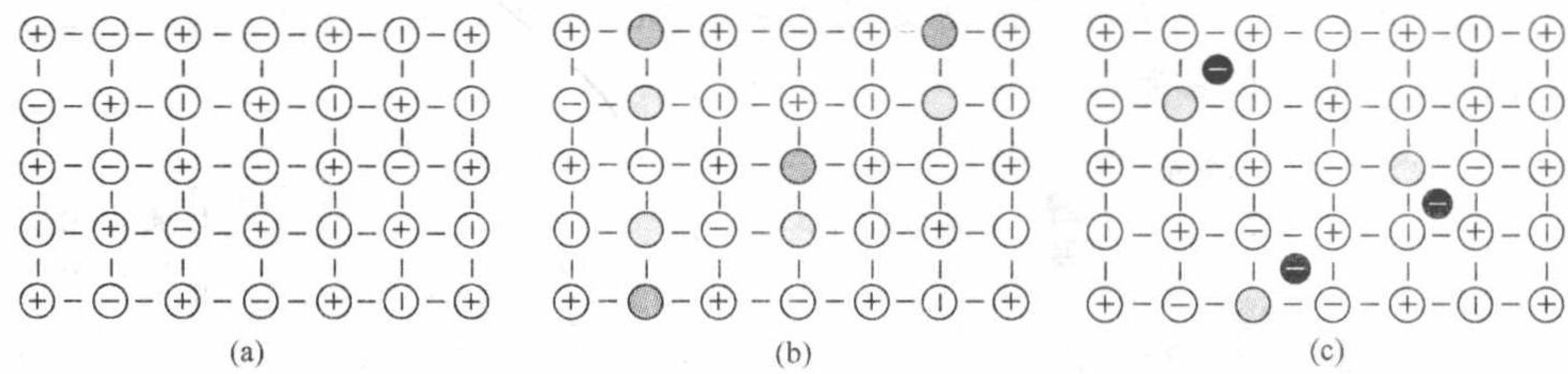


图 10.15 硅晶体中的空穴电子对

当 p 型和 n 型半导体结合在一起时，在两种半导体的交界面区域里会形成一个特殊的薄层，界面的 p 型一侧带负电，n 型一侧带正电. 这是由于 p 型半导体多空穴，n 型半导体多自由电子，出现了浓度差. n 区的电子会扩散到 p 区，p 区的空穴会扩散到 n 区，一旦扩散就形成了一个由 n 指向 p 的"内电场"，从而阻止扩散进行. 达到平衡后，就形成了这样一个特殊的薄层形成电势差，这就是 p-n 结.

当晶片受光后，p-n 结中，n 型半导体的空穴往 p 区移动，而 p 区中的电子往 n 型区移动，从而形成从 n 区到 p 区的电流. 然后在 p-n 结中形成电势差，这就形成了电源，如图 10.16 所示.

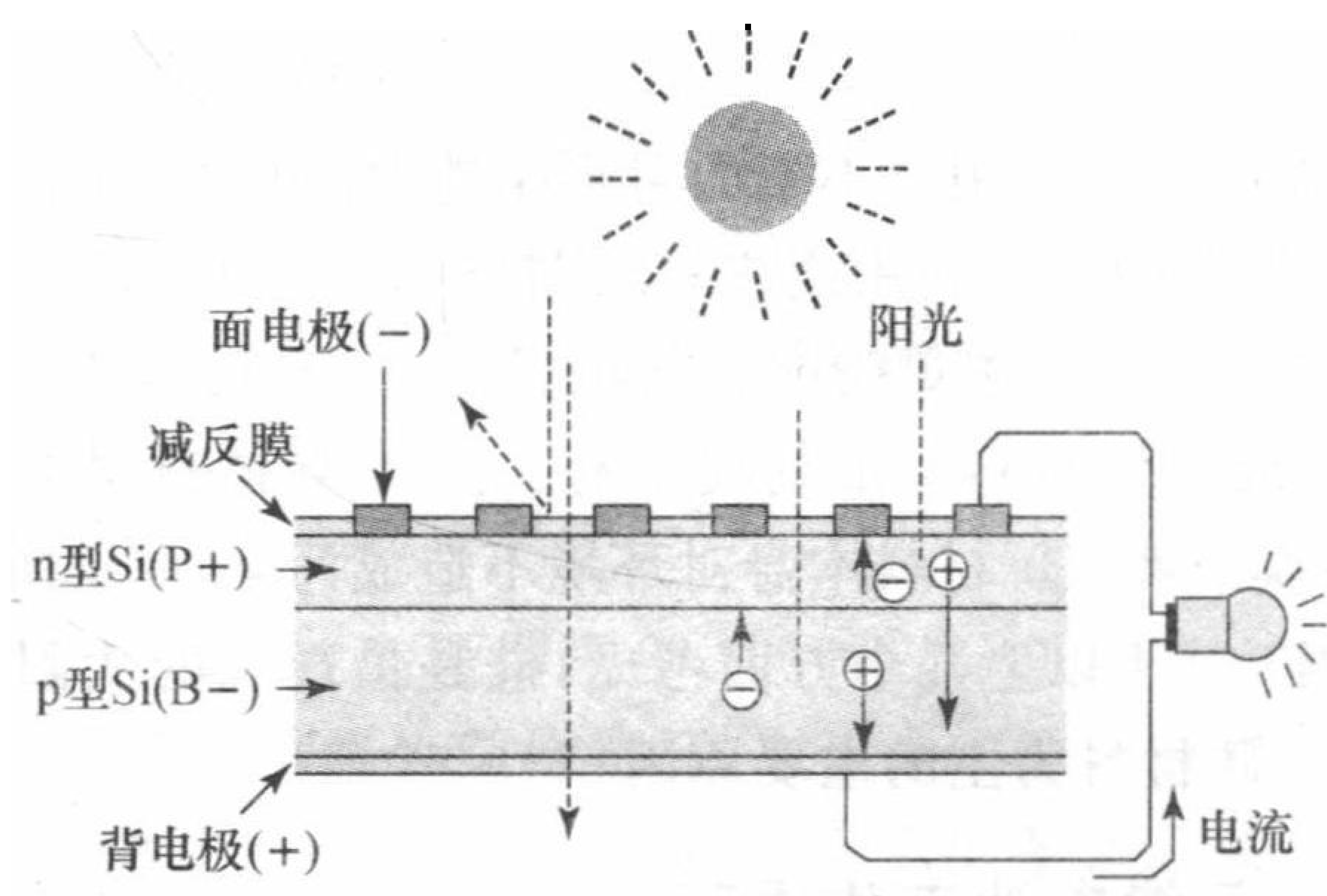


图 10.16 太阳能电池结构图

10.4.2 硅基太阳能电池

硅的光伏特性存在的独特优势，使其成为制作太阳能电池的良好材料. 硅基太阳能电池的研究和发展始终围绕着两个关键问题展开，光电转换效率和成本. 目前的硅基太阳能电池主要分为三类：单晶硅太阳能电池、多晶硅太阳能电池

和非晶硅薄膜太阳能电池.

单晶硅太阳能电池采用高纯的单晶硅为原料(纯度约为 9.99%),优点是光电转换效率比较高.典型的高效单晶硅太阳能电池是新南威尔士大学研制的钝化发射区背面局部扩散(passivated emitter rear locally diffused, PERL)太阳能电池[33](见图 10.17),这种电池表面具有倒金字塔织构、双层减反膜以及背反射结构,使电池拥有优异的陷光性能;利用氧化层钝化电池的正、背两面,增加了电池少子寿命;并且采用点接触代替原来的全覆盖式的背面铝合金接触,使 PERL 电池的转换效率高达 24.7%,接近理论值.但是,这种太阳能电池的生产需要消耗大量的高纯硅材料,而制造这些材料工艺复杂,电能消耗很大,成本较高.目前,超薄化电池发展趋势,使得实验室可以制备出 40 μm 厚的电池片,降低了单个太阳能电池对原材料的需求,或许可以使得单晶硅太阳能电池更具竞争力.

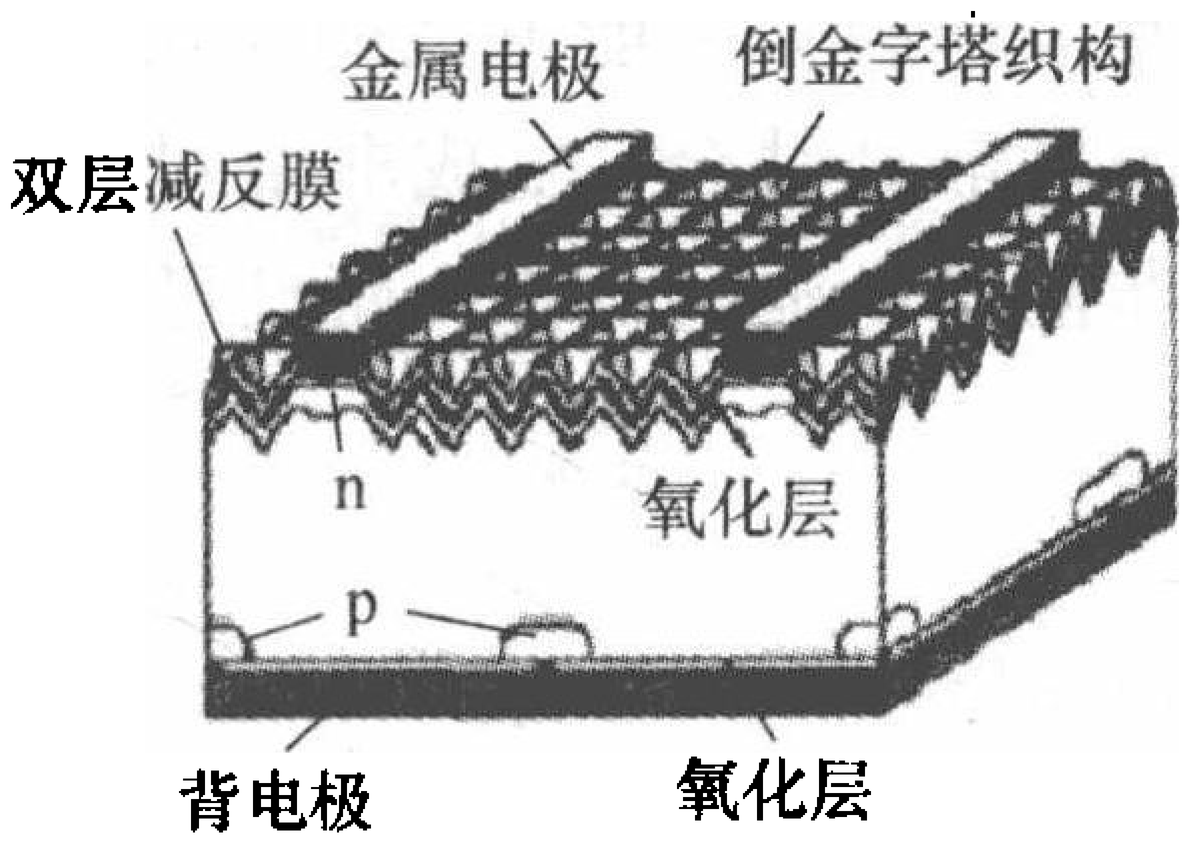


图 10.17　PERL 太阳能电池

多晶硅太阳能电池较单晶硅太阳能电池具有成本低的优势,但是光电转化效率较低,商业化多晶硅太阳能电池的效率约为 14%～18%.德国夫琅禾费(Fraunhofer)太阳能研究所制备的多晶硅太阳能电池(见图 10.18)的光电转换效率已经达到 20.3%[34],刷新了多晶硅太阳能电池转换效率的记录.这种电池不仅具有局部背表面场结构和用等离子体掩模法制备的表面织构,光学和电学性能良好,而且由于它采用了湿法氧化法而非传统的热氧化钝化电池后表面,在钝化效果和温度因素之间找到了一个合适的平衡点,既保证了钝化效果,又减少了温度对少子寿命的影响,使电池的性能得到最优化.

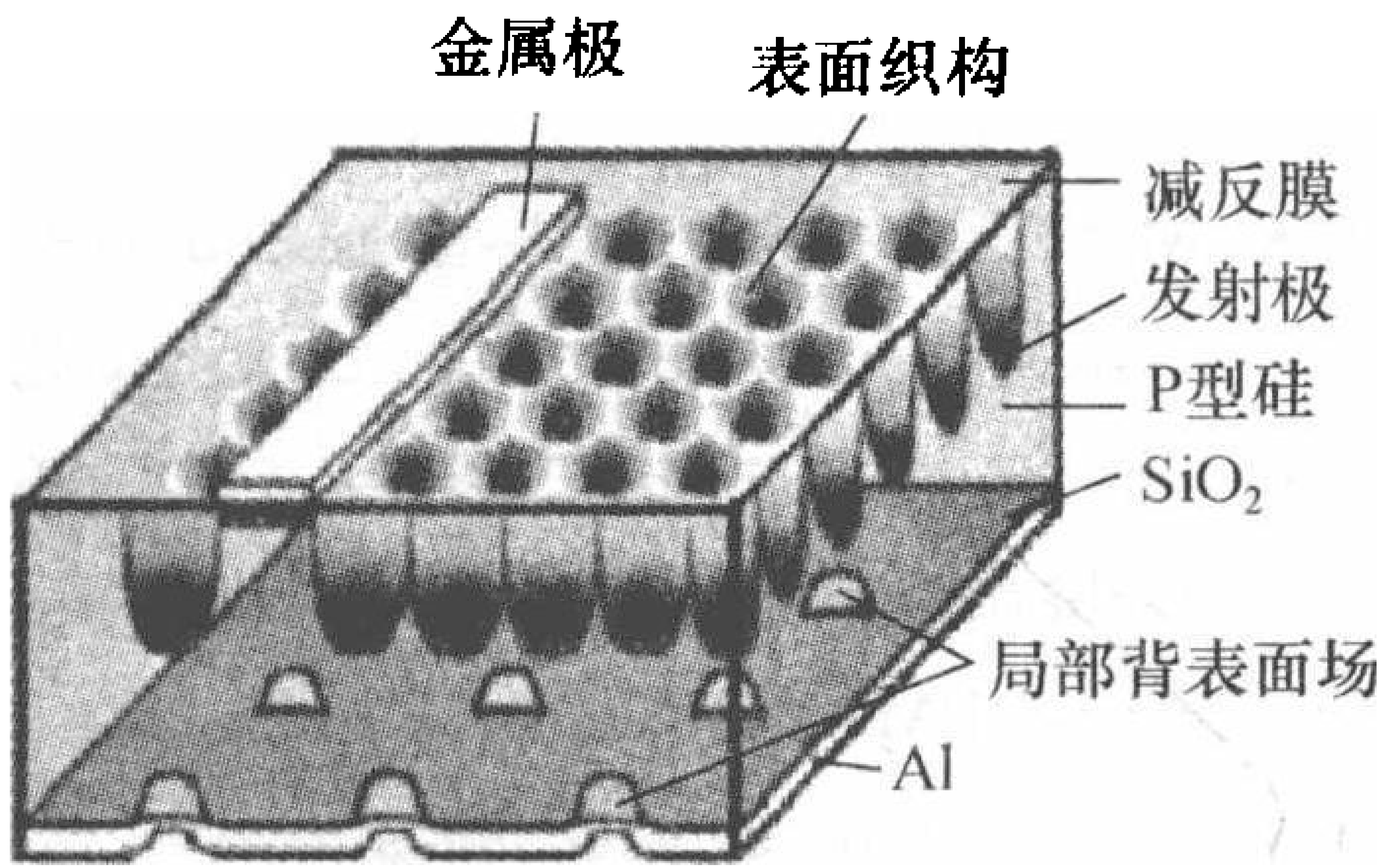


图 10.18　多晶硅太阳能电池(光电效率约为 20.3%)

非晶硅薄膜太阳能电池是一种以非晶硅化合物为基本组成的薄膜太阳能电

池.这种电池最大的特点是生产成本低,由于反应温度低,可在 200 ℃左右的温度下制造,因此可以在玻璃、不锈钢板、陶瓷板、柔性塑料片上沉积薄膜,易于大面积化生产,成本较低.但是由于非晶硅缺陷较多,制备的太阳能电池效率偏低,且其效率还会随着光照而衰减(Staebler-Wronski 效应,SW 效应)[35],导致非晶硅薄膜太阳能电池的应用受到限制.目前非晶硅薄膜电池研究的主要方向是与微晶硅结合,生成非晶硅/微晶硅异质结太阳能电池,这种电池不仅继承了非晶硅电池的优点,而且可以延缓非晶硅电池的效率随光照衰减的速度[36].

参考文献

[1] Intel Corporatin. Silicon photonics 50 G link. http://newsroom. intel. com/docs/DDC-1131, July 27th,2010.

[2] Liu A, et al. High-speed optical modulation based on carrier depletion in a silicon waveguide. Opt. Exp., 2007,15(2): 660.

[3] Liu A, et al. High-speed optical modulation based on carrier depletion in a silicon waveguide. Opt. Exp., 2007,15(2): 660.

[4] Yin T, et al. 31 GHz Ge n-i-p waveguide photodetectors on Silicon-on-Insulator substrate. Opt. Exp., 2007,15(21): 13965.

[5] Koch T, Koren U, Gnall R P, et al. GaInAs/GaInAsP multiple-quantum-well integrated heterodyne receiver. Electr. Lett., 1989,25(24): 1621.

[6] Takeuchi H, Kasaya K, Kondo Y, et al. Monolithic integrated coherent receiver on InP substrate. IEEE Photon. Tech. Lett., 1989,1(11): 398.

[7] Koch T L, Choa F S, Koren U, et al. Balanced operation of InGaAs-InGaAsP multiple quantum well integrated heterodyne receiver. IEEE Photo. Tech. Lett., 1990,2(8): 577.

[8] Deri R J, Sanada T, Yasuoka N, et al. Low-loss monolithic integration of balanced twin-photodetectors with a 3 dB waveguide coupler for coherent lightwave receivers. IEEE Photo. Tech. Lett., 1990,2(8): 581.

[9] Deri R J, Pennings E C M, Scherer A, et al. Ultracompact monolithic integration of balanced, polarization diversity photodetectors for coherent lightwave receivers. IEEE Photo. Tech. Lett., 1992,4(11): 1238.

[10] Bach H G, Matiss A, Leonhardt C C, et al. Monolithic 90° hybrid with balanced PIN photodiodes for 100 Gbit/s PM-QPSK receiver applications. Proc. Opt. Fiber Communication-includes post deadline papers, 2009.

[11] Doerr C R, Zhang L, Winzer P J, et al. Monolithic InP dual-polarization and dual-quadrature coherent receiver. IEEE Photo. Tech. Lett., 2011,23: 694.

[12] Doerr C R, et al. Monolithic polarization and phase diversity coherent receiver in silicon. J. Lightwave Technology, 2010,28(4): 520.

[13] Leung A, et al. A review of fiber-optic biosensors. Sensor Actuators B, 2007, 125: 688.

[14] Anker J N, et al. Biosensing with plasmonic nanosensors. Nature Materials, 2008, 7: 442.

[15] Fan X, et al. Sensitive optical biosensors for unlabeled targets: A review. Anal. Chim. Acta., 2008, 620: 8.

[16] Dandin M, et al. Optical filtering technologies for integrated fluorescence sensors. Lab on a Chip, 2007, 7: 955.

[17] Carlborg C F, et al. A packaged optical slot-waveguide ring resonator sensor array for multiplex label-free assays in labs-on-chips. Lab on a Chip, 2010, 10: 281.

[18] Cheben P, et al. Tilted fiber bragg grating sensor interrogation system using a high-resolution silicon-on-insulator arrayed waveguide grating. Opt. Lett., 2008, 33(22): 2647.

[19] Kribich K R, et al. Novel chemical sensor/biosensor platform based on optical multimode interference (MMI) couplers. Sensors and Actrators B, 2005, 107: 188.

[20] Dumais P, et al. Silicon-on-silicon optical sensor based on integrated waveguides and microchannels. IEEE Photo. Tech. Lett., 2005, 17(2): 441.

[21] Naoi Y, et al. Sensor for refractive index variation of an optical surface using a high-refractive-index waveguide. Opt. Eng., 2007, 46(10): 104601.

[22] Prasad T, et al. A photonic crystal sensor based on the superprism effect. Opt. mat., 2006, 29(1): 56.

[23] Hsu S H, et al. Design and analysis of Mach-Zehnder interferometer sensors based on dual strip antiresonant reflecting optical waveguide structures. Opt. Lett., 2005, 30 (21): 2897.

[24] Vos K D, et al. Silicon-on-Insulator microring resonator for sensitive and label-free biosensing. Opt. Exp., 2007, 15(12): 7610.

[25] Xu Q, et al. Silicon microring resonators with 1.5 μm radius. Opt. Exp., 2008, 16(6): 4309.

[26] Kiyat I, et al. High-Q silicon-on-insulator optical rib waveguide racetrack resonators. Opt. Exp., 2005, 13(6): 1900.

[27] Chao C Y, et al. Design and optimization of microring in biochemical sensing applications. J. Lightwave Technology, 2006, 24(3): 1395.

[28] Washburn A L, et al. Label-free quantitation of a cancer biomarker in complex media using silicon photonic microring resonators. Anal. Chem., 2009,81: 9499.

[29] Nitkowski A, et al. Cavity-enhanced on-chip absorption spectroscopy using microring resonators. Opt. Exp., 2008, 16(16): 11930.

[30] Wang J, et al. Highly sensitive Si nanowire-based optical sensor using a Mach-Zehnder interferometer coupled microring. Opt. Lett., 2010, 35(24): 4229.

[31] Xu D X, et al. Folded cavity SOI microring sensors for high sensitivity and real time measurement of biomolecular binding. Opt. Exp., 2008, 16(19): 15137.

[32] Zhu J, et al. On-chip single nanoparticle detection and sizing by mode splitting in an ultrahigh-Q microresonator. Nature Photonics, 2009, 4: 46.

[33] Chopra K L, et al. Thin-film solar cells. Prog. Photovolt: Res. Appl., 2004(12): 69.

[34] Schultz O, et al. Multi crystalline silicon solar cells exceeding 20% efficiency. Progress in Photovoltaics, 2004, 12(7): 553.

[35] Finger F, et al. Stability of microcrystalline silicon for thin film solar cell applications. IEE Proc. Circuits Devices Syst., 2003, 150(4): 300.

[36] Shah A V, et al. Thin-film silicon solar cell technology. Prog. Photovolt: Res. Appl., 2004(12): 113.